T.C.
GEBZE TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

ULTRASON GÖRÜNTÜLERİNDEN UZAM-ZAMANSAL DERİN AĞLARA DAYALI DİL KONTURU TESPİTİ

ENES ASLAN
YÜKSEK LİSANS TEZİ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

GEBZE
2019

T.C.
GEBZE TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

ULTRASON GÖRÜNTÜLERİNDEN UZAM-ZAMANSAL DERİN AĞLARA DAYALI DİL KONTURU TESPİTİ

ENES ASLAN
YÜKSEK LİSANS TEZİ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

DANIŞMANI
PROF. DR. YUSUF SİNAN AKGÜL

GEBZE
2019

T.R.

GEBZE TECHNICAL UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

TONGUE CONTOUR EXTRACTION FROM ULTRASOUND IMAGES USING SPATIOTEMPORAL DEEP NETWORKS

ENES ASLAN

A THESIS SUBMITTED FOR THE DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF COMPUTER ENGINEERING

THESIS SUPERVISOR

PROF. DR. YUSUF SINAN AKGUL

GEBZE

2019

GEBZE TEKNİK ÜNİVERSİTESİ	**YÜKSEK LİSANS JÜRİ ONAY FORMU**

GTÜ Fen Bilimleri Enstitüsü Yönetim Kurulu'nun 03/07/2019 tarih ve 2019/30 sayılı kararıyla oluşturulan jüri tarafından 16/07/2019 tarihinde tez savunma sınavı yapılan Enes ASLAN'ın tez çalışması Bilgisayar Mühendisliği Anabilim Dalında YÜKSEK LİSANS tezi olarak kabul edilmiştir.

JÜRİ

ÜYE (TEZ DANIŞMANI)	: Prof. Dr. Yusuf Sinan AKGÜL
ÜYE	: Doç. Dr. Erchan APTOULA
ÜYE	: Dr. Öğr. Üyesi Ayşe Betül OKTAY

ONAY

Gebze Teknik Üniversitesi Fen Bilimleri Enstitüsü Yönetim Kurulu'nun

...../...../..... tarih ve ...../..... sayılı kararı.

ÖZET

Ultrason teknolojisi, insanlar üzerinde zararlı etkilerinin bulunmaması, gerçek zamanlı veri sağlaması, görece ucuz olması gibi sebeplerden dolayı yaygın olarak kullanılmaktadır. Ultrason görüntülerinden dil konturu tespiti: Dilin modellenmesi, yabancı dil öğrenmede telaffuz eğitimi, dilin yapısal bozukluğundan meydana gelen kalıtsal hastalıkların tespiti, bazı üflemeli müzik enstrümanlarını çalma eğitimi gibi birçok alanda kullanılmaktadır. Bu tez çalışması kapsamında, derin öğrenme tabanlı, ultrason görüntülerinden dil konturu tespiti yapan iki farklı yöntem geliştirilmiştir.

İlk yöntemde, CNN, ultrason görüntüsü parçacıkları ile eğitilip oluşturulan model üzerinde tahmin yapılarak kontur tespit edilmiştir. Ultrason görüntüsü parçacıklarına, kontura uzaklığına göre sınıflar atanarak eğitim gerçekleştirilmiştir. Eğitilmiş model üzerinde kayan pencere yöntemi ile yapılan tahminler sayesinde her bir parçacığın dil konturu üzerinde olma ihtimali belirlenmiştir. Bu parçacıklar ve ihtimaller kullanılarak regresyon ile 3.dereceden bir eğri uydurulup dil konturu tespiti gerçekleştirilmiştir.

Önerilen ikinci yöntemde ise uzam-zamansal LSTM ağları eğitilerek dil konturu tespiti gerçekleştirilmiştir. Ultrason görüntüleri, polar koordinat sistemine çevrilerek görüntü matrisinin her bir kolonu, ultrason yansıtıcısından çıkan bir ışına denk gelecek şekilde LSTM'in zaman adımlarına verilerek eğitim gerçekleştirilmiştir. Uzamsal ve uzam-zamansal olarak iki tür LSTM ağı eğitilmiştir. Bu sayede uzamsal ve zamansal bilgiler birleştirilip video karelerinin arasındaki zaman bilgisi de kullanılarak başarım arttırılmıştır.

Çalışmada kullanılan veri kümesi, laboratuvar ortamında, ultrason cihazı kullanılarak, 5 farklı denek ile iki oturumda yapılan kayıtlardan oluşturulmuştur. Elde edilen başarım, literatürde şimdiye yapılmış derin öğrenme ve aktif kontur modeli tabanlı çalışmaların başarımına göre daha yüksektir. Sistem, başarımının yüksek olmasının yanında gerçek zamana yakın bir hız performansı da göstermektedir.

Anahtar Kelimeler: Yapay sinir ağları, Derin Öğrenme, Dil Konturu Tespiti, Ultrason, Evirişimsel Sinir Ağları (CNN), Kayan Pencere, Uzun-kısa Süreli Bellek (LSTM).

SUMMARY

Ultrasound technology is widely used for reasons such as lack of harmful effects on humans, providing real-time data, being relatively cheap. Tongue contour detection from ultrasound images; It is used in many areas such as language modeling, pronunciation training in foreign language learning, speech to text conversion, detection of hereditary diseases caused by tongue structural disorder and training to play some blown musical instruments. In this thesis, two different methods is developed that are based on deep learning and that perform tongue contour detection from ultrasound images.

In the first method, CNN was trained with ultrasound image parts and predicted on the created model to detect the contour. Training is carried out by assigning classes to ultrasound image parts, according to distance from tongue contour. The probability of each parts being on the tongue contour is determined by the sliding window method on the trained model. By using these parts and probabilities, a 3rd degree curve is fitted and a tongue contour is determined.

In the second method, the tongue contour is determined by training the spatio-temporal LSTM networks. Ultrasound images are transformed into the polar coordinate system and each column of the image matrix is trained by giving the time steps of the LSTM to coincide with a beam emitting from the ultrasound transducer. Two types of LSTM networks are trained as spatial and spatio-temporal. In this way, spatial and temporal information is combined and performance is improved by using the time information between video frames.

The data set used in the study is created from recordings made in two sessions with 5 different person, using ultrasound device, in laboratory environment. The final accuracy results are exceed the state of the art results. In addition to its high performance, the system also shows a speed performance close to real time.

Key Words: Deep Learning, Tongue Contour Extraction, Ultrasound, convolutional Neural Networks (CNN), Sliding Window, Long Short Term Memory (LSTM).

TEŞEKKÜR

Lisans ve yüksek lisans eğitimim boyunca yardımları, bilgileri ve tecrübeleri ile bana sürekli destek olan, tez çalışmam boyunca ilgisini, yönlendirmelerini esirgemeyen danışmanım Prof. Dr. Yusuf Sinan AKGÜL hocama, manevi destekleri ve bana olan inançları için aileme en içten teşekkürlerimi sunarım.

İÇİNDEKİLER

	Sayfa
ÖZET	v
SUMMARY	vi
TEŞEKKÜR	vii
İÇİNDEKİLER	viii
SİMGELER ve KISALTMALAR DİZİNİ	ix
ŞEKİLLER DİZİNİ	x
TABLOLAR DİZİNİ	xii
1. GİRİŞ	1
2. LİTERATÜR ARAŞTIRMASI	4
2.1. AKTİF KONTUR TABANLI ÇALIŞMALAR	5
2.1.1. KLASİK AKTİF KONTUR MODELLERİ	5
2.1.2. EDGETRAK	6
2.1.3. TONTRAK	6
2.2. DERİN ÖĞRENME TABANLI ÇALIŞMALAR	6
2.2.1. DBN	8
2.2.2. SU-NET	9
2.2.3. SEGNET & SU-NET	10
2.3. KARŞILAŞTIRMA VE SINIFLANDIRMA	10
3. ULTRASON GÖRÜNTÜSÜNÜ ANLAMA	12
4. VERİ KÜMESİ	14
5. CNN İLE DİL KONTURU TESPİTİ	15
5.1. SİSTEMİN EĞİTİMİ VE TESTİ	15
5.1.1. UZAMSAL RGB	15
5.1.2. UZAM-ZAMANSAL RGB (ZAMANSAL PARÇACIKLAR BİRLEŞİMİ)	16
5.1.3. UZAM-ZAMANSAL RGB (ZAMANSAL PARÇACIKLAR FARKI)	17
5.1.4. ÇOK ÖLÇEKLİ UZAM-ZAMANSAL RGB	18
5.2. PERFORMANS KARŞILAŞTIRMASI	19
5.3. 3. DERECEDEN EĞRİ UYDURMA	25

5.4. DENEYLER VE SONUÇLAR 25
5.5. ÖRNEK ÇIKTILAR 28
6. LSTM İLE DİL KONTURU TESPİTİ 29
6.1. KUTUPSAL KOORDİNAT SİSTEMİNE DÖNÜŞÜM 30
6.2. SİSTEMİN EĞİTİMİ VE TESTİ 31
6.3. RANSAC İLE EĞRİ UYDURMA 34
6.4. AĞ PARAMETRELERİ VE UZAM-ZAMANSAL AĞ ADEDİ 36
6.4.1. AĞ PARAMETRELERİ SEÇİMİ 36
6.4.2. UZAM-ZAMANSAL AĞ ADEDİ SEÇİMİ 37
6.5. DENEYLER VE SONUÇLAR 38
6.5.1. TEZ ÇALIŞMASINDAKİ AĞLARIN BİRBİRİ ÜZERLERİNDEKİ ETKİLERİ 39
6.5.2. LİTERATÜRDEKİ YÖNTEMLER İLE KARŞILAŞTIRMA 40
6.5.3. 5 KATLAMALI ÇAPRAZ DOĞRULAMA 40
6.5.4. REFERANS NOKTASI DENEYİ 41
6.6. ÖRNEK ÇIKTILAR 43
7. SONUÇLAR 44

KAYNAKLAR 46
ÖZGEÇMİŞ 53
EKLER 54

SİMGELER ve KISALTMALAR DİZİNİ

Simgeler ve Kısaltmalar		**Açıklamalar**
CNN	:	Convolutional Neural Network
RNN	:	Recurrent Neural Network
DBN	:	Deep Belief Network
RBM	:	Restricted Boltzmann Machine
LSTM	:	Long Short Term Memory
RGB	:	Red Green Blue
RAM	:	Random Access Memory
CPU	:	Central Processing Unit
MSD	:	Mean Sum of Distance
RANSAC	:	Random Sample Consensus

ŞEKİLLER DİZİNİ

Şekil No:		Sayfa
2.1:	Ultrasonla dil konturu tespiti yapan literatür çalışmalarının gruplanması.	4
2.2:	Ultrason görüntüsü üzerinde çalışan örnek bir CNN.	7
2.3:	RBM ağının katmanları.	8
2.4:	U-Net'in sadeleştirilmiş hali olan sU-net ağı.	9
2.5:	Dil konturu tespiti yapan çalışmaların kullandıkları modele göre sınıflandırılması	10
3.1:	Ultrason modları ve ultrason görüntüsünün elde ediliş şekilleri.	12
4.1:	(a)'da ultrason cihazından kayıt alma, (b)'de ise ultrason videosundan ayrıştırılmış örnek görüntü gösterilmiştir.	14
5.1:	Sistemde kullanılan AlexNet ağının, ultrason görüntülerinde çalışması için özelleştirilmiş sürümü.	15
5.2:	Uzam-zamansal parçacıklar birleşimi ve parçacıklar farkı ile veri kümesi oluşturulması.	17
5.3:	Sistem çıktılarının sınıfları ile oluşturulmuş ısı haritası. Kırmızı renge yaklaştıkça dil konturu olma ihtimali artmakta, mavi renge yaklaştıkça azalmaktadır.	19
5.4:	Çok ölçekli uzam-zamansal RGB veri kümesi oluşturulması.	19
5.5:	Uzamsal model.	20
5.6:	Uzam-zamansal RGB (zamansal parçacıklar birleşimi) model.	21
5.7:	Uzam-zamansal RGB (zamansal parçacıklar farkı) model.	21
5.8:	Çok ölçekli uzam-zamansal RGB model.	22
5.9:	CNN ile dil konturu tespiti sisteminin eğitim aşamaları.	23
5.10:	CNN ile dil konturu tespiti sisteminin test aşamaları.	24
5.11:	CNN ile dil konturu tespiti yapan sistemin ölçüm metriği.	26
5.12:	CNN ile dil konturu tespiti sonuçları. 1. satırda ham ultrason görüntüleri, 2.satırda el ile işaretlenmiş dil konturları, 3.satırda ise sistemin bulduğu dil konturları gösterilmiştir.	22
6.1:	Kutupsal ve dikdörtgensel koordinat sisteminde ultrason görüntüleri.	30

6.2: LSTM ile dil konturu tespiti yapan sistemin eğitim ve test aşamaları. 31
6.3: LSTM'in bulduğu noktalar ve bu noktalara RANSAC ile uydurulmuş eğriler gösterilmiştir. 35
6.4: İki kontur üzerindeki nokta sayısı eşit olmasa bile MSD metriği ile ölçüm gerçekleştirilebilir. 38
6.5: Örnek U-Net çıktıları. 42

TABLOLAR DİZİNİ

Tablo No:		Sayfa
2.1:	Literatürdeki dil konturu tespiti yapan sistemlerin sınıflandırılması ve başarımlarının karşılaştırılması.	11
4.1:	Deneklerden elde edilen görüntü sayıları.	14
5.1:	CNN ile kontur bulan sistemin sonuçları.	27
6.1:	Ağ parametreleri ve hata oranları.	36
6.2:	Uzamsal ve zamansal ağ sayısı ve hata oranları.	37
6.3:	LSTM ile dil konturu bulan sistemde kullanılan ağların birbirlerinin başarımları üzerindeki etkileri.	39
6.4:	LSTM ile dil konturu bulma sonuçlarının literatürdeki yöntemlerin sonuçları ile karşılaştırılması.	40
6.5:	5 katlamalı çapraz doğrulama deneyleri sonuçları.	41

1. GİRİŞ

Ultrason teknolojisi, zararsız olması, organların yapısının gerçek zamanlı olarak izlenebilmesini sağlaması ve görece ucuz olmasından dolayı oldukça yaygın olarak kullanılmaktadır. Geçmişten günümüze dilin ultrason görüntüsü üzerinde çalışmak, araştırmacılar için oldukça ilgi çekici bir alan olmuştur. İnsan dili hareketleri, konuşma esnasında dışarıdan görülemediği için ultrason cihazı ile görüntülenir. Farklı kelimeler için dilin aldığı şekiller izlenir ve böylece dil modellenebilir. Ultrason cihazı, dilin modellenmesinde, görüntüleme [21] ve takip [22] işlemleri için kullanılır. Ultrason cihazı, konuşma esnasında, dilin esnek ve hızlı hareketlerini [57], [58] 2 boyutlu dünyada takip edebilmek için yüksek çerçeve oranı ile donatılmıştır (60Hz ile 120Hz arası). 3 boyutlu dünyada kayıt yapan ultrason cihazları daha düşük çerçeve oranlarına sahiptir [59], bu sebepten dolayı görüntü işleme alanında çok sık kullanılmazlar.

Ultrasondan elde edilen dil konturu görüntüleri birçok alanda kullanılmaktadır. Son yıllarda yapılan çalışmalar, ikinci dil öğrenme eğitimleri [52], [23] üzerine yoğunlaşmıştır. Görsel geri bildirim, insanların hareketlerini düzenlemelerinde önemli bir araçtır [54], [55]. Bu motor öğrenme prensibi olarak da bilinir [56]. Müzik gibi alanlarda, eğitmenler, öğrencileri değerlendirmek için sıklıkla bu yöntemi kullanırlar [54].

Konuşma, insanlar arasında iletişimi sağlayan en önemli kanallardan biridir. Konuşma esnasında telaffuz hataları, yanlış anlaşılmalara ve iletişim kopukluklarına sebep olabilmektedir. Özellikle yabancı dil öğrenen insanlarda, sıklıkla, telaffuz problemleri yaşanmaktadır. Kelimelerin düzgün bir şekilde söylenmesi için dilin farklı şekiller alması gerekmektedir. İnsanların konuşma esnasında bu şekilleri görmesi ve kendilerini eğitmesi mümkün değildir. Dil, ultrason ile görüntülenerek, aldığı şekiller takip edilir. Telaffuz eğitiminde, lisanı sonradan öğrenen kişiye, ana lisanlı kişinin dil ultrasonu görüntüsü ve kendi dilinin anlık ultrason görüntüsü gösterilerek kelimeleri doğru telaffuz etmesi için eğitilebilir [7]. Literatürde, dilin ultrason görüntülerinden lisan eğitimi için yapılmış oldukça fazla sayıda çalışma mevcuttur [22]-[26], [60]-[63]. Dilin ultrason görüntüsünün kullanıldığı bir diğer alan ise konuşma terapisidir. Bazı insanlar, konuşma esnasında, çeşitli [27] sorunlar yaşayabilmektedirler. Dilin ultrason ile görselleştirilmesi, bu tip sorunlar yaşayan insanlara, konuşurken dilini nasıl hareket ettirmesi gerektiği konusunda yardımcı olmaktadır [51]. Dil ultrasonunun bir başka

kullanım alanı ise medikal alandır. Dilin yapısında meydana gelen bozukluklardan büyük veya küçük bazı hastalıkların tespiti yapılmaktadır. Bunlara kanser veya bazı alerjik kabarcıklar örnek olarak gösterilebilir.

Literatürde dil konturu tespiti alanında yapılmış birçok çalışma bulunmaktadır. Öncelerde kontur tespiti aktif kontur [35] tabanlı çalışmalar ile yapılmaktaydı [3]-[8]. Son yıllarda derin öğrenmenin popülerleşmesi ve klasik yöntemlere göre yüksek başarım elde edilmesi ile derin öğrenme tabanlı yöntemlere rastlanmaktadır [1], [7], [9], [21]. Bölüm 2'de literatürdeki bu alanda yapılmış çalışmalar detaylı olarak anlatılacaktır.

Literatürdeki dil konturu bulmak için yapılmış çalışmaların hepsi kontur bulma problemini klasik piksel sınıflandırma [36] problemi olarak ele almaktadır. Oysa ki ultrason görüntüsündeki dil, bir konturdur. Ultrason görüntüsü doğası gereği oldukça fazla gürültü içermektedir. Sınıflandırma yapan yöntemlerin konturun modelinden haberi olmadığı için bu gürültüleri de konturmuş gibi sınıflandırabilmektedir. Bunun yanında ultrason görüntüleri video çerçevelerinden elde edilmektedir. Kontur, ardışık çerçeveler arasında çok fazla pozisyon değişikliğine uğramamaktadır. Bu sebepten dolayı çerçeveler arası zamansal bilgi, kontur tespit etme problemi için kullanılırsa başarımın artacağı bir gerçektir. Literatürdeki çalışmalar incelendiğinde, uzamsal bilginin yanında zamansal bilgiyi de kullanan derin öğrenme tabanlı bir çalışmaya rastlanmamaktadır. Bu tez çalışmasında, derin öğrenme tabanlı iki yöntem ile dil konturu tespiti gerçekleştirilmiştir.

İlk yöntemde, CNN [31] ağları, ultrason görüntüsü parçacıkları ile eğitilmiştir. Konturları el ile işaretlenmiş ultrason görüntüleri, rastgele parçacıklara bölünmüştür. Her bir parçacığa, kontura olan Öklid uzaklığına [34] göre sınıf atanarak ağ eğitilmiştir. Bu parçacıklar oluşturulurken, bir önceki ve bir sonraki görüntülerde aynı merkez noktasına denk gelen parçacıklar da alınıp, boş bir RGB görüntünün R, G ve B kanallarına verilerek zamansal bilgi de eğitime dahil edilmiştir. Elde edilen model test edilirken, kayan pencere yöntemi kullanılmıştır. Her bir parçacık için bir sınıf tahmin edilmiş ve elde edilen sınıflardan, kontura belirli bir eşik değeri kadar yakın olanları kullanılarak, regresyon [33] ile 3. dereceden bir eğri uydurulup kontur tespiti gerçekleştirilmiştir.

İkinci yöntemde ise uzam-zamansal LSTM [32] ağlar kullanılarak dil konturu tespiti yapılmıştır. Bu yöntemde aynı zamanda ultrason görüntüsünün doğasına uygun bir eğitim gerçekleştirilmiştir. Yani sistem yalnızca dilin ultrason görüntülerinden

kontur tespiti için özelleşmiştir, literatürdeki diğer yöntemlerde olduğu gibi piksel sınıflandırması yapmamaktadır. İlk olarak ultrason görüntüleri kutupsal koordinat sistemine çevrilmiştir. Bu sayede ultrason görüntü matrisindeki her bir kolon ultrason yansıtıcısından çıkan bir dalgaya denk gelmiştir. Aynı zamanda her bir kolonda en fazla bir adet kontur noktası bulunacağı veya hiç bulunmayacağı anlamına da gelmektedir. Böylece arama uzayı ciddi ölçüde küçültülerek, sistem yalnızca dil konturu modeline uygun bir çıktı verecek şekle getirilmiştir.

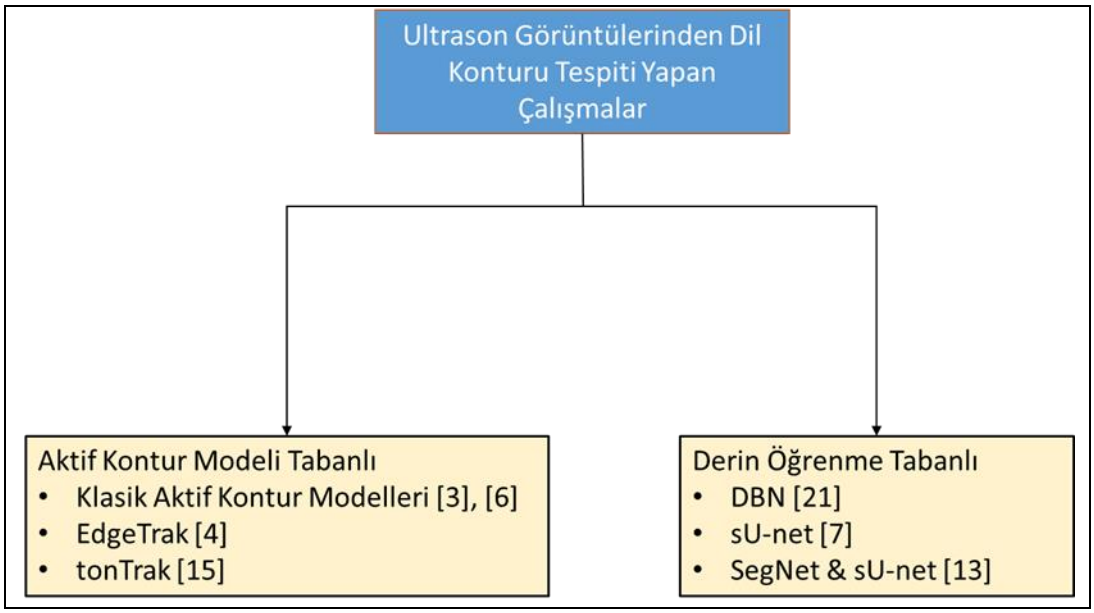
Uzamsal ve zamansal olarak iki tür LSTM ağı eğitilmiştir. Eğitim esnasında bu ağların çıktıları birbirlerine girdi olarak verilerek uzamsal ve zamansal bilgiler birleştirilmiştir. Uygulanan bu yöntem ile literatürdeki çalışmalara göre daha yüksek bir başarım elde edilmiştir.

Her iki sistemin eğitim ve test aşamalarında kullanılan veri kümeleri laboratuvar ortamında 5 denek kullanılarak 2 oturumda yapılan ultrason kayıtlarından elde edilmiştir.

Tezin 5. bölümünde CNN ile yapılan çalışma, 6. bölümünde uzam-zamansal LSTM ağları ile yapılan çalışma detaylı olarak anlatılmıştır. Son bölümde ise sonuçlar ve çıkarımlar yer almaktadır.

2. LİTERATÜR ARAŞTIRMASI

Literatürdeki ultrason görüntülerinde dil konturu tespiti yapan uygulamalara bakıldığında, geçmiş yıllarda makine öğrenmesi olmadan, görüntü özellikleri kullanılarak yapılmış çalışmalara rastlanmaktadır. Bu çalışmalar çoğunlukla aktif kontur kullanarak dil konturu tespiti yapmaktadır [3]-[6], [8]. Son yıllarda grafik kartlarının gelişmesi ile bilgisayarla görme alanındaki çalışmalarda, genelde derin öğrenme kullanılmaya başlanmıştır. Derin öğrenme, yüksek başarım oranı, hızlı çalışması, manuel veri çıkartma yerine otomatik olarak eğitim verisi çıkartması gibi sebeplerden dolayı tercih edilmektedir. Ultrason görüntülerinden dil konturu tespiti yapan çalışmalarda kullanılan aktif kontur da yerini derin öğrenmeye bırakmıştır. [21], [7], [9], [21] derin öğrenme ile dil konturu yapan çalışmalara örnek olarak gösterilebilir.



Şekil 2.1: Ultrasonla dil konturu tespiti yapan literatür çalışmalarının gruplanması.

2.1. Aktif Kontur Tabanlı Çalışmalar

Aktif kontur modeli [35], yılan olarak da bilinmektedir. 2 boyutlu, gürültülü görüntülerden, nesnelerin ana hatlarını çıkartmak için kullanılır. Bilgisayarla görme alanında oldukça sık kullanılan bir yöntemdir. Nesne takip etme [76], şekil tanıma [77], sınıflandırma [36], kenar tespit etme [78] gibi alanlarda yaygın olarak kullanılmaktadır. Aktif kontur, oluşturulan bir enerji fonksiyonunu minimize ederek bulunmak istenen nesneyi veya konturu çıkartır. Bu fonksiyon iç ve dış enerjilerin toplamından oluşmaktadır. İç enerji, görüntü üzerinde bir noktanın tek başına hedef kontur veya nesneye ne kadar uygun olduğunu ifade ederken dış enerji, noktanın komşuları ile olan ilişkisini ifade eder. Bu enerji fonksiyonuna göre, noktalar kümesi yer değiştirilerek hedef nesne veya kontura doğru yinelemeler ile yaklaştırılır. Aktif konturun dezavantajı manuel olarak başlatma gerektirmesi ve bazı durumlarda konturun hedeften uzaklaşarak gürültülere veya istenmeyen bölgelere takılmasıdır.

2.1.1. Klasik Aktif Kontur Modelleri

Aktif kontur modeli ile gerçekleştirilmiş çalışmalardan [3]'de sürekli olmayan bir aktif kontur modeli kullanılmıştır. Aktif konturlar temel olarak aynı işi yapsa da problemin yapısına göre formülde bazı değişiklikler yapılarak kullanılabilmektedir. Bu çalışmada gürültüsü fazla olan görüntüler üzerinde çalışması için özelleştirilmiş bir aktif kontur modeli kullanılmıştır. Kullanılan ultrason görüntüleri, denek başını ve ultrason yansıtıcısını destekleyici bir sistem aracılığı ile kayıt edilmiştir. Bu sistemin kullanılma sebebi konuşma esnasında ultrason yansıtıcısının pozisyonunu bozulmasını engellemektir. Sistem tasarımında genelden özele bir yol izlenmiştir, buna göre ilk olarak görüntülerdeki piksellerin bir alt kümesi alınıp bunlar üzerinde kontur tespit edilmiş, sonrasında ise bu alt küme büyütülerek yeniden kontur tespiti yapılmıştır. Bu yaklaşımla gürültüye karşı gürbüzlük sağlanmıştır. Bir başka aktif kontur tabanlı dil konturu takibi yapan çalışma [6]'dır. Bu çalışmada dil konturu, konturun hareketinin zaman içerisindeki değişimi izlenerek takip edilmiştir. Görüntüler arasında gürültü veya kayıt esnasında ultrason yansıtıcısının hareket etmesi sonucu meydana gelen kopmalar için otomatik olarak ilklendirme yapan bir yöntem geliştirilerek kopma sorununa da çözüm üretilmiştir. Sistem, dinamik programlama ile eniyilenmiştir.

2.1.2. EdgeTrak

Bir diğer aktif kontur modeli tabanlı dil konturu takibi yapan çalışma [4]'tür. Çalışmada EdgeTrak adı verilen bir kontur takip sistemi geliştirilmiştir. Sistem, kullanıcıdan girdi olarak bir görüntü için kontur üzerinde noktalar kümesi alır, bu noktaları kullanarak B-spline [30] interpolasyon yöntemi ile o görüntü için dil konturu yakınsaması yapar, dinamik programlama ile de kontur yüzeyini bulur. Her görüntü için, bir önceki görüntüde bulunan kontur noktaları kullanılarak, aktif kontur modeli ile ilklendirme yapılır. Dinamik programlama ile de eniyileme gerçekleştirilir. Çalışmadaki aktif kontur modeli, [3]'te kullanılan gradyan tabanlı modelin parlaklık kısıtı eklenmiş halidir. Böylece kontur yüzeyindeki piksellere benzerlik gösteren gürültülerin, elenmesi sağlanmıştır. Bu kısıta göre, iki nokta arasındaki parlaklık farkı sıfırdan küçükse, bu noktalar için belirli bir miktarda ceza uygulanır, bu sayede modelin başarımı iyileştirilir.

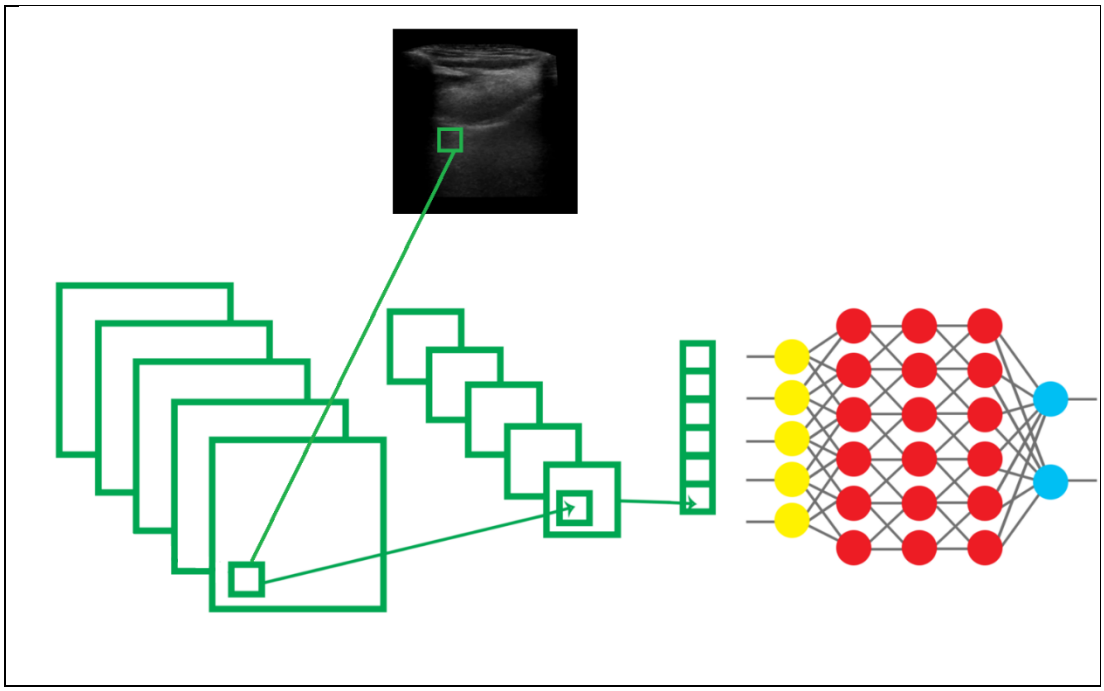
2.1.3. tonTrak

Aktif kontur modeli kullanan çalışmalardan [15]'in motivasyonu, tez çalışmasının LSTM ile dil konturu bulan çalışmasının motivasyonuna benzemektedir. TonTrak adı verilen sistem, sadece dil konturu bulmak için özelleşmiş bir sistemdir. Bu sisteme göre, ultrason cihazından yayılan dalgaların, dil konturunu bir kez kestiği varsayılır. Bu varsayım ile dil konturu için arama uzayı küçültülmüş olur. Çalışmanın temel motivasyonu, ultrason yansıtıcısının görüntü oluşturma işlemindeki, yansıtıcıdan çıkan dalgalar ile dil konturunun kesişim sayısı limitine dayanmaktadır. Kontur takibi için, diğer çalışmalardan farklı olarak, 3 boyutlu uzayda çalışan bir aktif kontur modeli kullanılmıştır. Model, zamansan ve uzamsal bileşenlerden oluşturulmuştur. Eniyileme için [16]-[19] çalışmalarında bahsedilen dinamik programlama yöntemi kullanılarak, [17], [18, [20] çalışmalarında yaşanan yerel minimum probleminin önüne geçilmiştir.

2.2. Derin Öğrenme Tabanlı Çalışmalar

Son yıllarda, ultrason görüntülerinden dil konturu yapan çalışmalar incelendiğinde, çoğunlukla derin öğrenme ile geliştirilen sistemler olduğu görülmektedir. Bilgisayar donanımlarının güçlenmesi, çok çekirdekli, güçlü grafik

kartlarının üretilmesi ve yaygın olarak kullanılmaya başlanması, bilgisayar ile hesaplama hızını oldukça arttırmıştır. Bu gelişmeler sonucunda, eskilerden beri bilinen fakat hesaplama maliyetinin yüksek olmasından ve bunu yapacak bilgisayarların olmamasından dolayı pratikte kullanılamayan derin sinir ağları popüler bir hale gelmiştir. Bu ağlar temel olarak CNN ve RNN olmak üzere iki guruba ayrılmaktadırlar.



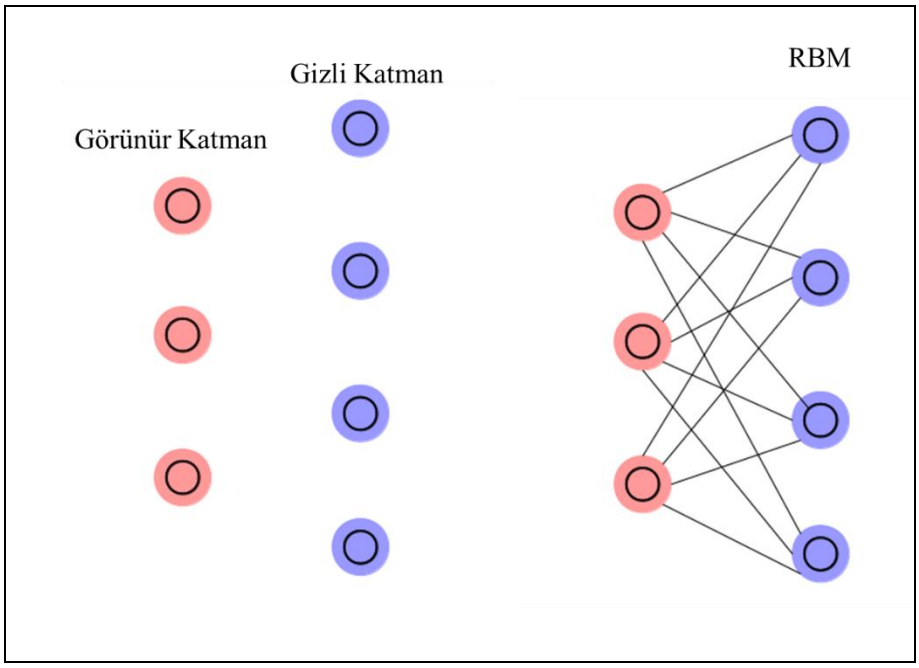
Şekil 2.2: Ultrason görüntüsü üzerinde çalışan örnek bir CNN.

CNN'ler girdi olarak görüntü alıp, görüntü içerisindeki nesneleri veya birbiri ile ilişkili piksel gruplarını ayırt etmek için, ağırlık ataması yapar. Klasik makine öğrenme yöntemlerine göre daha az ön işleme gerektirir. Bu tür derin ağlar, çok basit ve çeşitli ilkel filtreleri, evrişim ile, görüntü alanlarına uygular ve görüntü içerisindeki cisimleri birbirinden ayırır (Şekil 2.2).

Diğer derin ağlar ise RNN'lerdir. CNN'ler girdileri anlık olarak değerlendirir. RNN'ler ise girdileri sadece anlık değil, önceki girdilere bağlı olarak da değerlendirir. RNN'de girdi verilerinin yanında önceki çıktıları gösteren içerik birimleri de ağa etki ederler. t-1 anındaki girdi için verilen karar, t anında verilecek olan kararı da etkiler. Yani bu tür ağlarda girdiler şimdiki ve önceki bilgilerin birleştirilmesi ile çıktı üretirler. Literatürde, bu bölümde bahsedilen her iki derin öğrenme tekniği ile dil konturu yapan çalışmalar [1], [21] mevcuttur.

2.2.1. DBN

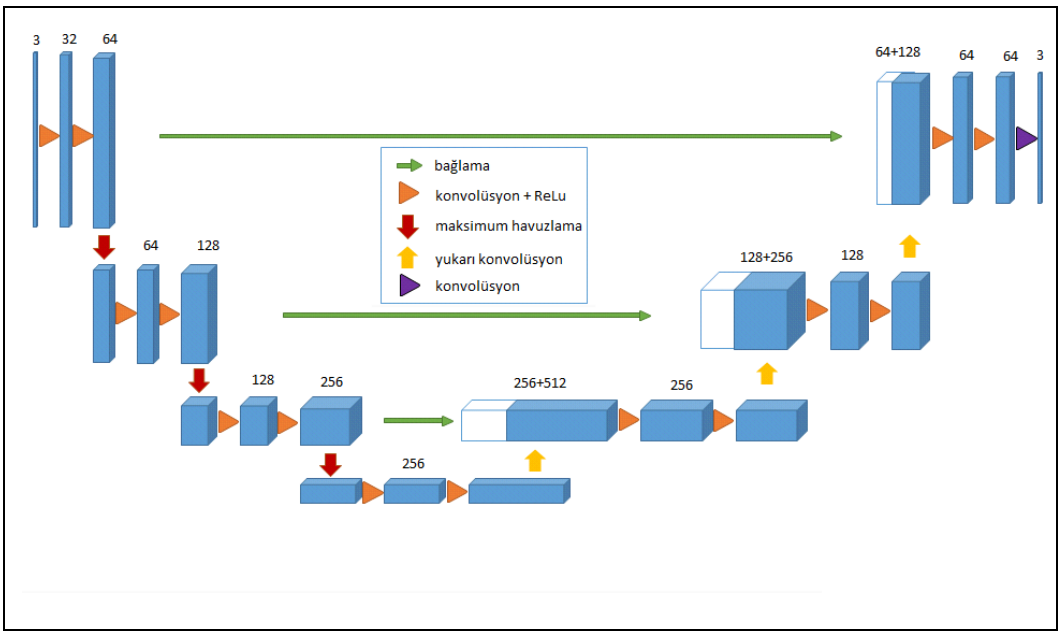
Derin öğrenme ile dil konturu tespiti yapan çalışmalardan [21]'de DBN kullanılmıştır. Ağ, birden fazla RBM'nin yığın olarak birleştirilmesi ile oluşturulmuştur. RBM, girdi seti üzerinde, olasılık dağılımlarını öğrenebilen, rastlantısal bir yapay sinir ağıdır. Her düğüm bir nöronu temsil eder ve aynı zamanda hesaplamaların yapıldığı yerdir. Düğümler, katmanlar arasında birbirine bağlanır fakat aynı katmanın iki düğümü birbirine bağlı değildir. Yani, katmanlar arası iletişim bulunmaz, bu yüzden Kısıtlı Boltzmann Makinesi olarak adlandırılır. Her düğümde girdiler hesaplanır ve hesaplanan girdinin bir sonraki düğüme iletimine ilişkin rastgele kararlar verir. RBM'nin ilk katmanı, görünür veya giriş katmanı, ikinci katmanı ise gizli katman olarak adlandırılır (Şekil 2.3). Sistem otomatik kodlayıcı ve kod çözücü olarak iki parçadan oluşacak şekilde tasarlanmıştır. Eğitim aşamasında, konturu işaretlenmiş ultrason görüntülerini alır. Test aşamasında ise ham ultrason görüntüsünü alıp konturu tespit edilmiş görüntüyü üretir.



Şekil 2.3: RBM ağının katmanları.

2.2.2. sU-net

Dil konturu tespiti yapan çalışmalardan [7]'de sU-net adı verilen evrişimsel bir derin ağ kullanılmıştır. sU-net, U-net ağlarının sadeleştirilmiş halidir. U-net [28] ağları medikal görüntü işleme alanında oldukça sık kullanılan evrişimsel derin ağlardır. Eşit sayıda yukarı örnekleme ve aşağı örnekleme katmanlarının birleştirilmesi ile oluşturulurlar. Çalışmada, U-net ağı üzerinde bazı sadeleştirmeler yapılmıştır. 23 olan



Şekil 2.4: U-Net'in sadeleştirilmiş hali olan sU-net ağı.

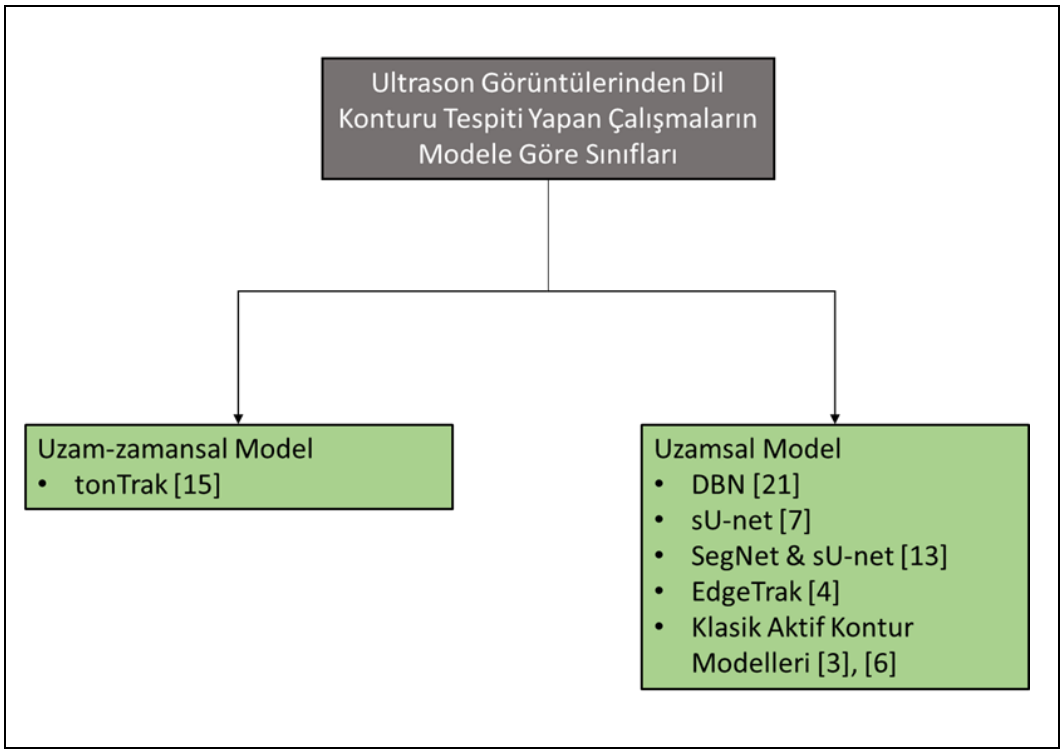
evrişim katmanı sayısı 14'e düşürülmüştür. Bunun yanında, her katmandaki aktivasyon fonksiyonu sayıları da azaltılmıştır. Bu değişikliklerle, eğitim esnasındaki binlerce parametre elenmiştir. Değişikliklerin temel motivasyonu ultrason görüntülerinin videodan alınmış olması sebebi ile ardışık görüntülerin birbirine çok benzemesidir. Şekil 2.4'te sU-net ağının tasarımı gösterilmiştir. Yapılan bu işlemler sonucunda sistem, gerçek zamanlı uygulamalarda çalışacak bir hıza getirilmiştir. Ayrıca daha az görüntü ile de eğitilip başarılı sonuçlar üretmesi sağlanmıştır.

Sistemin ürettiği kontur çıktılarında meydana gelen bozukluk ve kopuklukları engellemek amacı ile genişleme, erozyon gibi temek morfolojik [37] görüntü işleme yöntemleri kullanılmıştır.

2.2.3. SegNet & sU-net

[13]'te gerçekleştirilen uygulamada, gerçek zamanlı olarak dil konturu takibi yapılmaktadır. Uygulamanın motivasyonu, yabancı dil öğrenen kişilerdeki telaffuz bozukluklarının giderilmesi için, kişilerin eğitilmesini sağlamaktır. Bu amaç için, eğitilecek kişinin dilinin konturunu gerçek zamanlı olarak takip eden, bunun yanında ana dil konuşan kişilerden alınmış ultrason görüntülerini gösteren bir ara yüz tasarlanmıştır. Çalışmada derin ağ mimarisi olarak [7]'de geliştirilen sU-net ağı SegNet [38] ağı ile birleştirilerek kullanılmıştır. Sistemin gerçek zamanlı olarak çalışabilmesi için, eğitim esnasında, ultrason görüntüleri 128x128 boyutundan 34x34 boyutuna küçültülmüştür. Kontur tespit edildikten sonra, görüntüler tekrardan 128x128 boyutuna getirilmiştir. Sistem, Windows işletim sistemi, 7 çekirdekli 3.4 GHz frekansında CPU ve 16GB RAM ile saniyede yaklaşık olarak 74 kare işleyebilmektedir.

2.3. Karşılaştırma ve Sınıflandırma



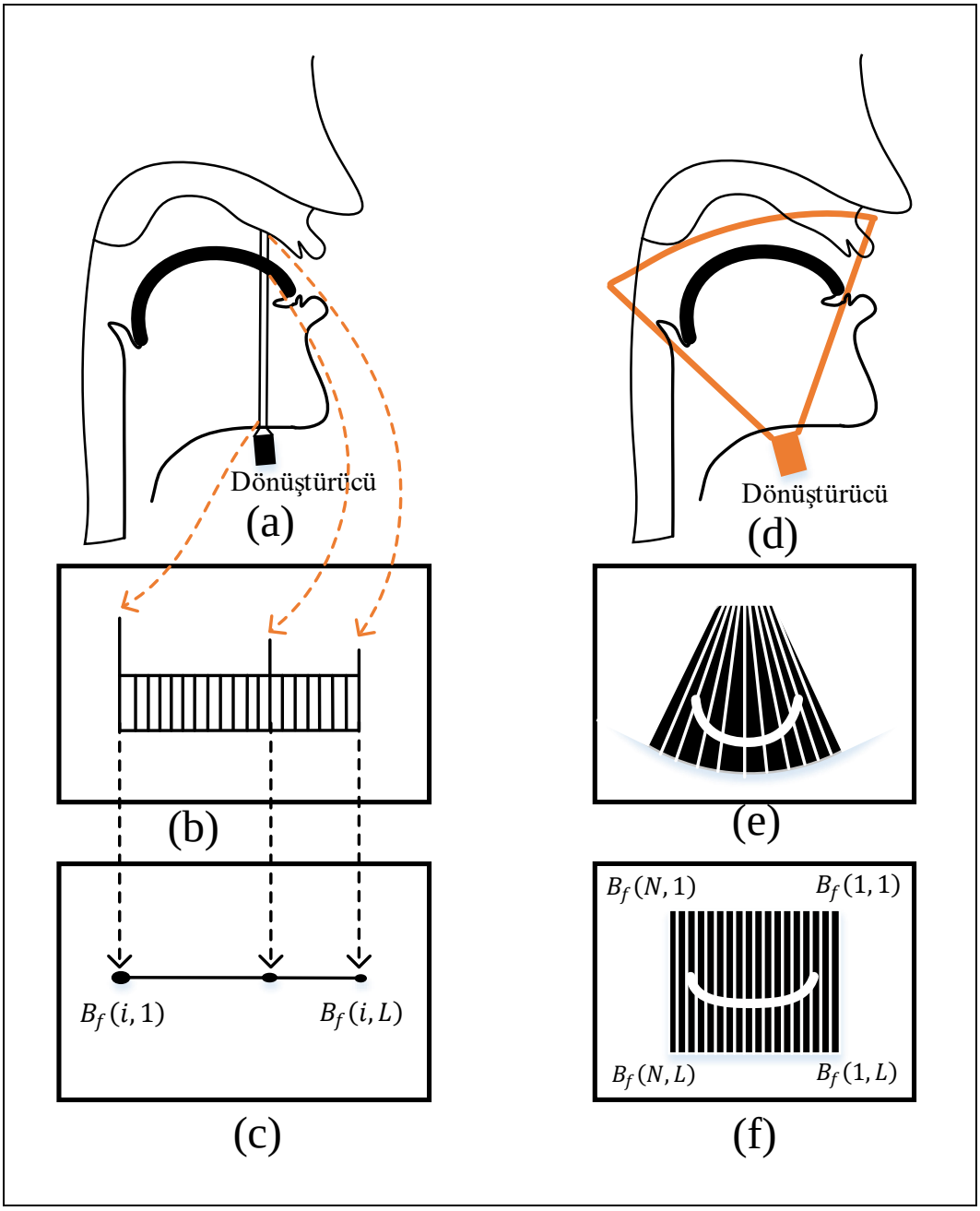
Şekil 2.5: Dil konturu tespiti yapan çalışmaların kullandıkları modele göre sınıflandırılması.

Literatürdeki yöntemler incelendiğinde, yalnızca [15]'in uzam-zamansal model kullandığı görülmektedir, fakat çalışma, aktif kontur modeli ile gerçekleştirilmiştir. Günümüzde, aktif konturlar gibi, görüntü özelliklerini herhangi bir öğrenme olmaksızın yorumlayarak sonuç üreten sistemler, yerini derin öğrenme tabanlı sistemlere bırakmıştır. Literatürdeki dil konturu tespiti yapan çalışmaların başarımları karşılaştırıldığında, en düşük hata oranının [7]'de olduğu gözlenmektedir. Uzam-zamansal model kullanılmamış olmasına karşın derin öğrenmenin gücü kullanılarak yüksek başarım elde edilmiştir.

Tablo 2.1: Literatürdeki dil konturu tespiti yapan sistemlerin sınıflandırılması ve başarımlarının karşılaştırılması.

Adı	Yılı	Model	Derin Öğrenme / Aktif Kontur	Başarım (mm)
Automatic contour tracking in ultrasound images [4]	2005	Uzam-zamansal	Aktif Kontur	0.5
A Task-Specific Contour Tracker for Ultrasound [15]	2000	Uzamsal		-
Deep belief networks for real-time extraction of tongue contours from ultrasound during speech [21]	2010		Derin Öğrenme	1.0
Tongue Contour Segmentation using Deep Learning [7]	2018			0.91
Real-time Automatic Tongue Contour Tracking in Ultrasound Video for Guided Pronunciation Training [13]	2019			0.44

3. ULTRASON GÖRÜNTÜSÜNÜ ANLAMA



Şekil 3.1: Ultrason modları ve ultrason görüntüsünün elde ediliş şekilleri.

Tez çalışmasında uygulanan yöntemin anlaşılması için, ultrason görüntüsünün nasıl elde edildiği, hangi çeşit ultrason görüntüleri olduğu konularında bilgi sahibi olmak gerekmektedir. Ultrason görüntüsü, yansıtıcıdan [64] yansıyan [66] ses dalgalarının [65] geri dönüş süreleri ölçülerek oluşturulmaktadır. Ultrason görüntüsünü oluşturmanın farklı çeşitleri bulunmaktadır. Şekil 3.1 (a)'da tek bir

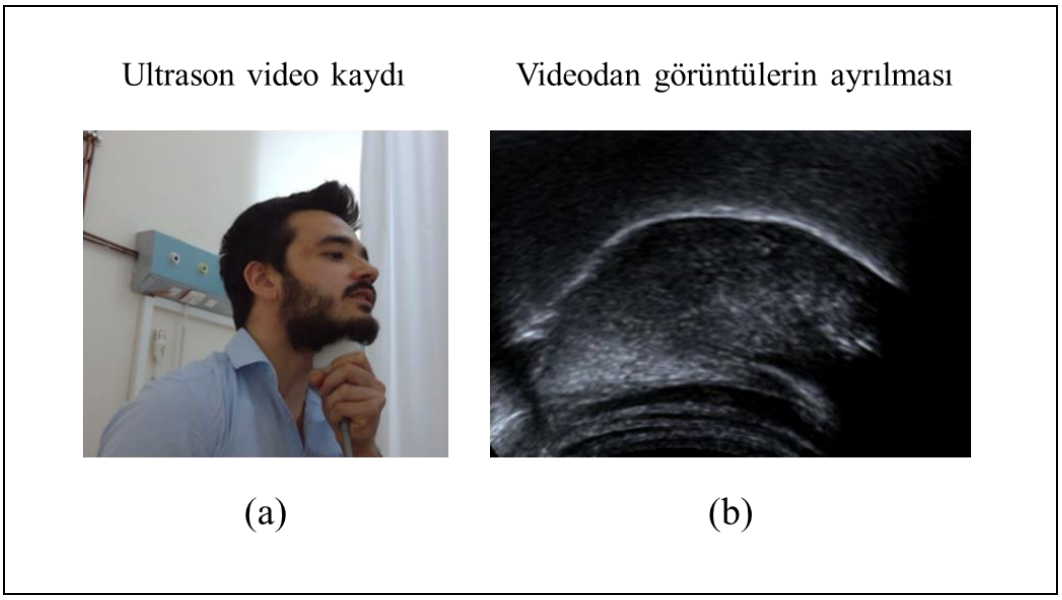
ultrason dönüştürücüsünden dile doğru giden bir dalga, Şekil 3.1 (b)'de ise genlik modu ultrason görüntüsü gösterilmektedir. Bu gösterim, t anında gönderilen dalganın dönüşlerinin genliklerinden oluşmaktadır. Şekil 3.1 (c)'de tek bir dalga için parlaklık modu ultrason görüntüsü gösterilmektedir. Bu gösterimde ise geri yansıyan dalgaların genlikleri yerine görüntü üzerindeki parlaklıkları kullanılmaktadır. Birden fazla ultrason dalgasının geri yansımaları ölçülüp birleştirilerek 2 boyutlu parlaklık modu görüntüsü elde edilir (Şekil 3.1 (e)). Bu tez çalışmasında 2 boyutlu parlaklık modu görüntüleri kullanılmıştır.

Ultrason görüntüsü, yansıtıcıdan çıkan dalgaların geri dönüşüne göre oluşturulmaktadır, yansıyan dalgalar organların içerisinde dağılarak gelmektedir bu da görüntünün yüksek seviyede gürültü içermesine neden olmaktadır. Bu gürültüye benek gürültüsü adı verilmektedir [70]. Son yıllarda geliştirilmiş ultrason cihazları, gürültü sorununu bir ölçüde azaltsa da [67] tam anlamı ile bitirememiştir. Gürültüden dolayı ultrason görüntülerinde kopukluklar [69] meydana gelmektedir. Ultrason görüntüleri üzerinde çalışan sistemlerin bu sorunun üstesinden gelecek şekilde tasarlanması gerekmektedir [68]-[71].

Benek gürültüsü rastgele oluşmaktadır [72], [73]. Ultrason görüntüsünün kalitesini arttırmak için bu gürültüyü azaltmak bir seçenektir fakat gürültü, dalgaların organlardan yansıması sonucu oluştuğu için organlar hakkında bilgi de içermektedir. Bir başka seçenek bu gürültü dağılımını bir desene sokup anlamlandırmaktır. Literatürde bu konu üzerine birçok çalışma mevcuttur [74], [11], [39], [53], [75].

4. VERİ KÜMESİ

Sistemin eğitim ve test aşamalarında, laboratuvar ortamında utlrason cihazından kaydedilmiş videolardan çıkartılan görüntüler kullanılmıştır (Şekil 4.1). Video kayıtları Toshiba Aplio500 model ultrason cihazı ile 5 denek kullanılarak 2 oturumda gerçekleştirilmiştir. Deneklerin tümünün cinsiyeti erkektir ve konuşma ile ilgili herhangi bir hastalıkları bulunmamaktadır. Tez çalışmasında kullanılan görüntüler tüm deneklere, "O güzel sesimi göstermek için burada hazırım." cümlesinin okutulduğu video kayıtlarından alınmıştır. Deneklerden elde edilen görüntü sayıları Tablo 4.1'de gösterilmiştir. Tüm görüntülerin çözünürlükleri 960x720'dir. Görüntüler üzerindeki dil konturları, geliştirilen bir araç ile eğitim ve doğrulama adımlarında kullanılmak üzere işaretlenmişlerdir.



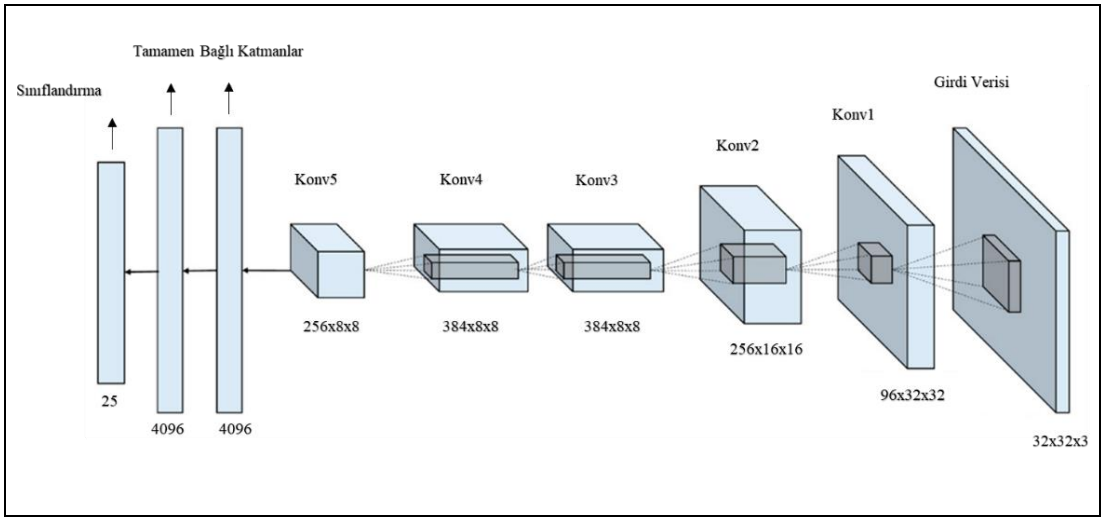
Şekil 4.1: (a)'da ultrason cihazından kayıt alma, (b)'de ise ultrason videosundan ayrıştırılmış örnek görüntü gösterilmiştir.

Tablo 4.1: Deneklerden elde edilen görüntü sayıları.

Denek Numarası	1. Oturum Görüntü Sayısı	2. Oturum Görüntü Sayısı
1	44	56
2	131	88
3	121	117
4	114	95
5	134	129

5. CNN İLE DİL KONTURU TESPİTİ

Bu çalışmada obje tanıma için tasarlanmış CNN tabanlı AlexNET [31] ağı kullanılmıştır. Ağ orijinalinde 256x256 boyutunda girdileri kabul etmektedir. Çalışmada kullanılan görüntü parçacıkları 32x32 ve 64x64 boyutlarında olduğu için ağın bu boyutlarda girdi alacak şekilde düzenlenmiş iki sürümü kullanılmıştır [9]. Şekil 5.1'de ağın 32x32 boyutunda girdileri alacak şekilde özelleştirilmiş sürümünün tasarımı gösterilmiştir. Benzer şekilde de 64x64 girdileri alacak ağ tasarlanmıştır.



Şekil 5.1: Sistemde kullanılan AlexNet ağının, ultrason görüntülerinde çalışması için özelleştirilmiş sürümü.

5.1. Sistemin Eğitimi ve Testi

5.1.1. Uzamsal RGB

Çalışma, eğitim ve test olmak üzere iki aşamadan oluşmaktadır. Sistemde kullanılacak görüntüler, ultrason videolarından ayrıştırılarak elde edilmiştir. Ayrıştırılan görüntüler eğitim ve test kümesi olmak üzere ikiye bölünmüştür. Görüntülerdeki dil konturları, geliştirilen bir araç ile eğitim ve test adımlarında kullanılmak üzere işaretlenmiştir. Böylece konturun üzerindeki her bir noktanın x, y koordinatları elde edilmiştir. Denklem 5.1'de u görüntü üzerindeki noktaları, v kontur üzerindeki noktaları, D ise bu noktalar arasındaki Öklid uzaklığını ifade etmektedir. Buna göre, merkez noktasının kontura uzaklığı $0 \leq D \leq 25$ olacak şekilde, görüntü üzerinden tesadüfi olarak alınan 32x32'lik parçacıklara, 0 ile 25

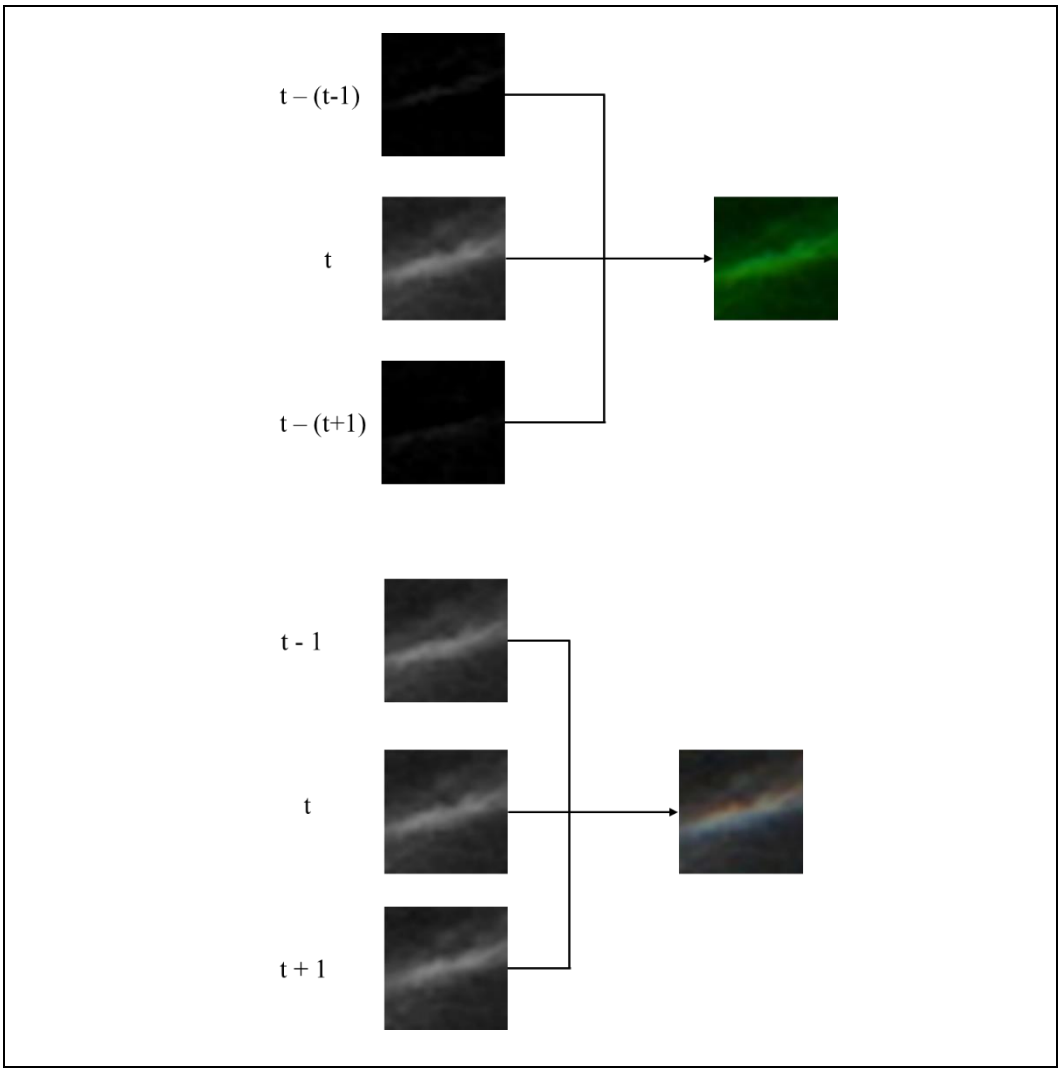
arasında bir sınıf atanmıştır. Elde edilen parçacıklar ve sınıflar kullanılarak eğitim ve test işlemleri gerçekleştirilmiştir.

$$D = \sqrt{(u_x - v_x)^2 + (u_y - v_y)^2} \tag{5.1}$$

5.1.2. Uzam-zamansal RGB (Zamansal Parçacıklar Birleşimi)

Sistemde kullanılan görüntüler videodan alındığı için zamana göre sıralıdır. Kontura uzaklığına göre sınıfları belirlenen görüntü parçacıklarının merkez koordinatlarının bir önceki ve bir sonraki görüntülerdeki aynı koordinatlara denk gelen parçacıklar da alınıp bir eğitim verisi parçacığı için Şekil 5.2'de görüldüğü gibi 3 adet parçacık elde edilmiştir. Bu parçacıklar 32x32 boyutunda üç kanallı bir görüntünün RGB kanallarına eklenerek renkli ve tek bir görüntü parçacığı oluşturulmuştur. Tüm parçacıklar için bu işlem gerçekleştirilmiştir. Böylece görüntülerin videodan alınmış olmasının avantajı kullanılarak zaman bilgisi de eğitimin içerisine dahil edilmiştir.

Sınıfları bilinen görüntü parçacıkları kullanılarak ağ Şekil 5.9'da gösterildiği gibi eğitilmiştir. Test aşamasında kayan pencere yöntemi kullanılmıştır. Test görüntülerinden 32x32'lik parçacıklar alınarak, eğitim aşamasında olduğu gibi bir önceki ve bir sonraki görüntülerdeki karşılıkları da alınıp 32x32 boyutunda boş bir RGB görüntüsünün R, G ve B kanallarına eklenmiştir. Şekil 5.10'da gösterildiği gibi eğitilmiş model aracılığı ile tahminler yapılıp her bir parçacığın sınıfı elde edilmiştir. Şekil 5.3'te elde edilen sınıflardan oluşturulmuş ısı haritası görülmektedir.



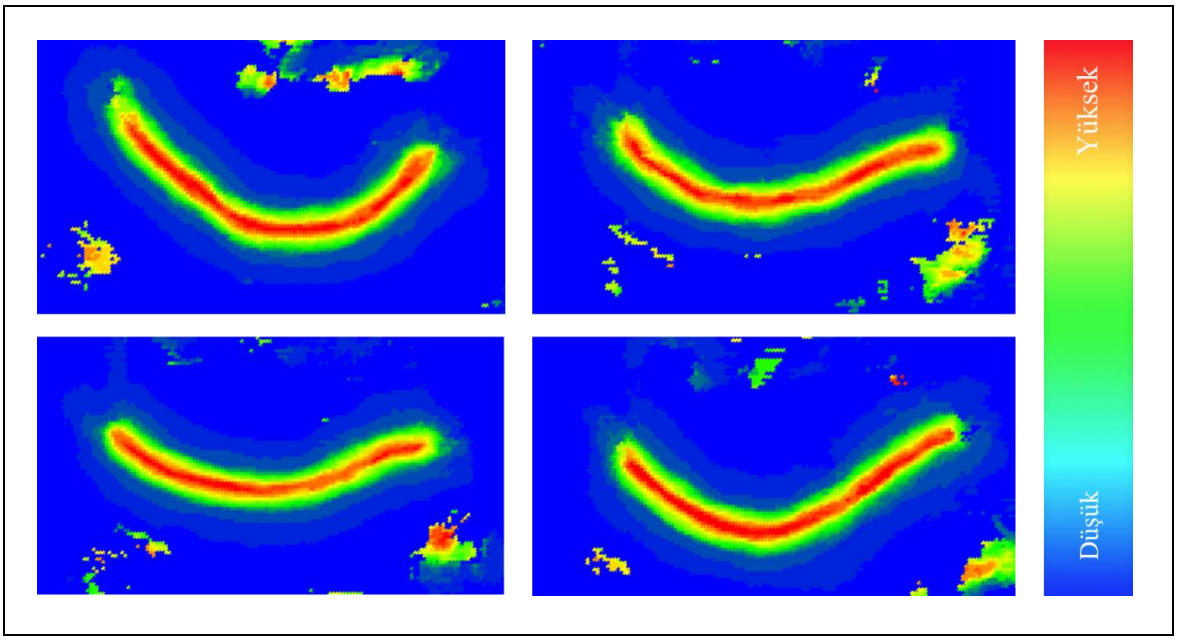
Şekil 5.2: Uzam-zamansal parçacıklar birleşimi ve parçacıklar farkı ile veri kümesi oluşturulması.

5.1.3. Uzam-zamansal RGB (Zamansal Parçacıklar Farkı)

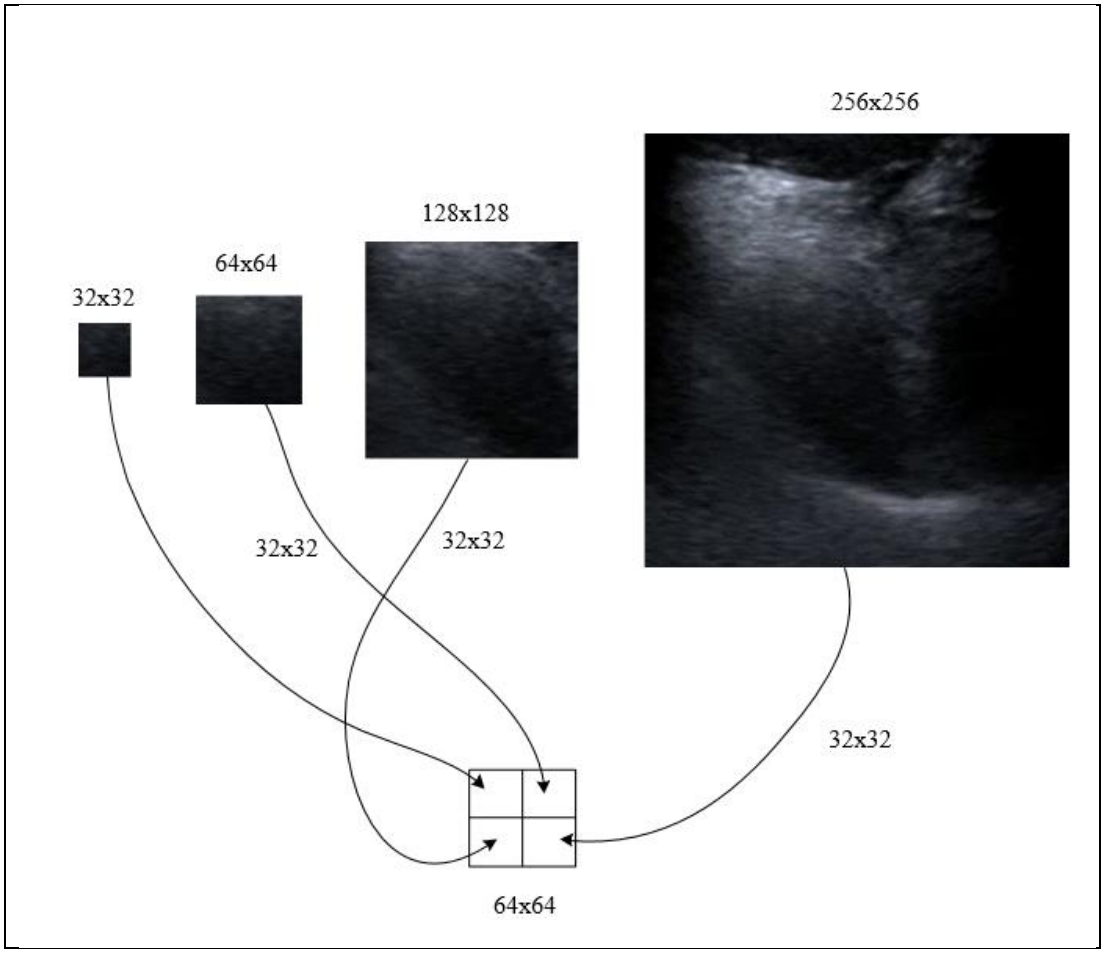
Bu bölümde 4.1.1'de uygulanan yöntemin çok benzeri uygulanmıştır. 32x32'lik parçacıklar ilgili görüntüden ve merkezi aynı noktaya denk gelen bir sonraki ve bir önceki görüntülerden de alınmıştır. Elde edilen 3 adet 32x32'lik parçacıklar ilgili görüntü parçacığı ile bir önceki görüntüdeki aynı merkeze denk gelen parçacığın farkı ve bir sonraki görüntüde aynı merkeze denk gelen parçacığın farkı alınmıştır. İlgili parçacık ve 2 fark parçacığı RGB görüntünün R, G ve B kanallarına eklenmiştir. 4.1.1'de parçacıklar olduğu gibi eklenirken bu çalışmada farkları alınarak eklenmiştir.

5.1.4. Çok Ölçekli Uzam-zamansal RGB

Önceki bölümlerde uygulanan teknikler parçacık tabanlı olduğu için 32x32'lik parçalara göre öğrenmekte ve tahmin yapmaktadır. Dolayısı ile ultrason görüntüsünü bir bütün halinde değerlendirememektedir. Bu çalışmada da önceki bölümlerde anlatıldığı gibi bir önceki ve bir sonraki kareler eğitim için RGB görüntüye eklenmek sureti ile kullanılmıştır. Önceki çalışmalardan farklı olarak merkez olarak seçilen bir nokta için, nokta merkezi olacak şekilde 32x32 boyutunda, 64x64 boyutunda, 128x128 boyutunda ve 256x256 boyutunda toplamda 4 adet parçacık alınmıştır. Tüm parçacıklar 32x32 boyutuna getirilerek 4 adet 32x32 boyutunda parçacık elde edilmiştir. Bu parçacıklar 64x64'lük boş bir görüntünün sol üstüne 32x32'lik parçacık, sağ üstüne 64x64'ten 32x32'ye küçültülmüş parçacık, sol altına 128x128'den 32x32'ye küçültülmüş parçacık ve sağ altına da 256x256'dan 32x32'ye küçültülmüş parçacık eklenmiştir. Böylece elimizde 64x64'lük ama 256x260 kadar bilgi içeren bir parçacık olmuştur (Şekil 5.4). Önceki yöntemlere göre bir parçacık ultrason görüntüsü ile çok daha fazla bilgi içermektedir. Bu değişim ile birlikte tasarım olarak 256x256 boyutunda girdiler alan AlexNet ağının da 64x64 boyutunda girdiler alacak şekilde güncelleştirilmesi gerekmiştir. Ağın giriş bölümü değiştirilip 64x64'e uygun hale getirilerek eğitim gerçekleştirilmiştir.



Şekil 5.3: Sistem çıktılarının sınıfları ile oluşturulmuş ısı haritası. Kırmızı renge yaklaştıkça dil konturu olma ihtimali artmakta, mavi renge yaklaştıkça azalmaktadır.

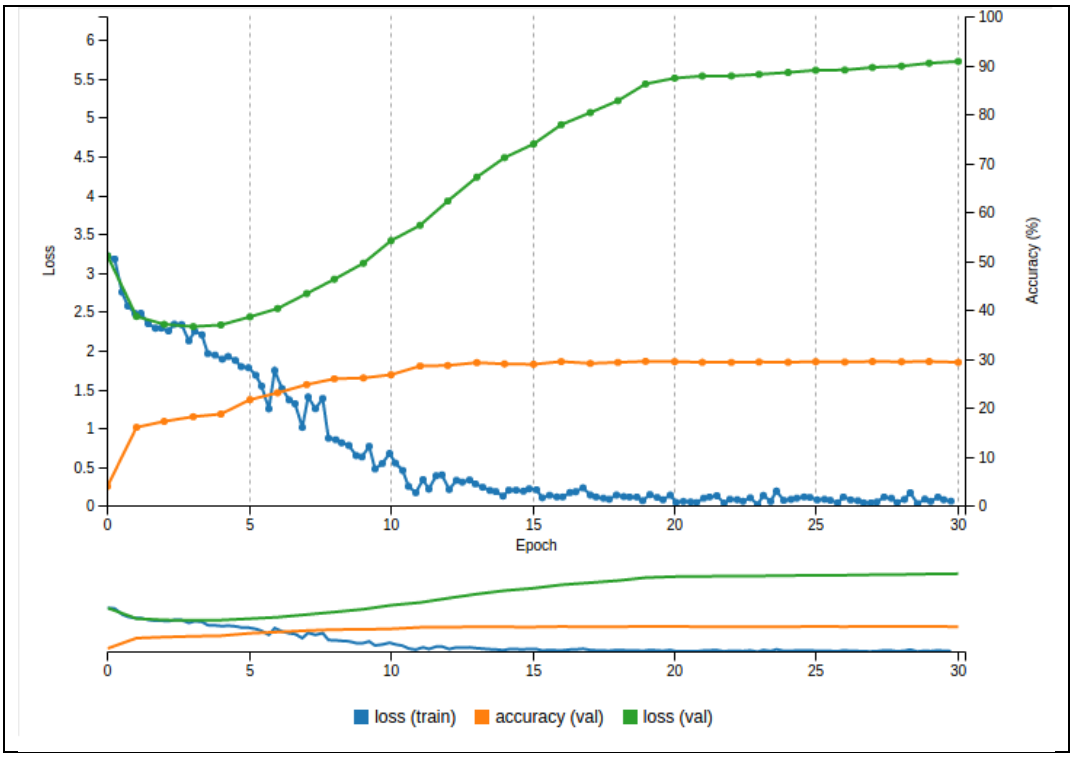


Şekil 5.4: Çok ölçekli uzam-zamansal RGB veri kümesi oluşturulması.

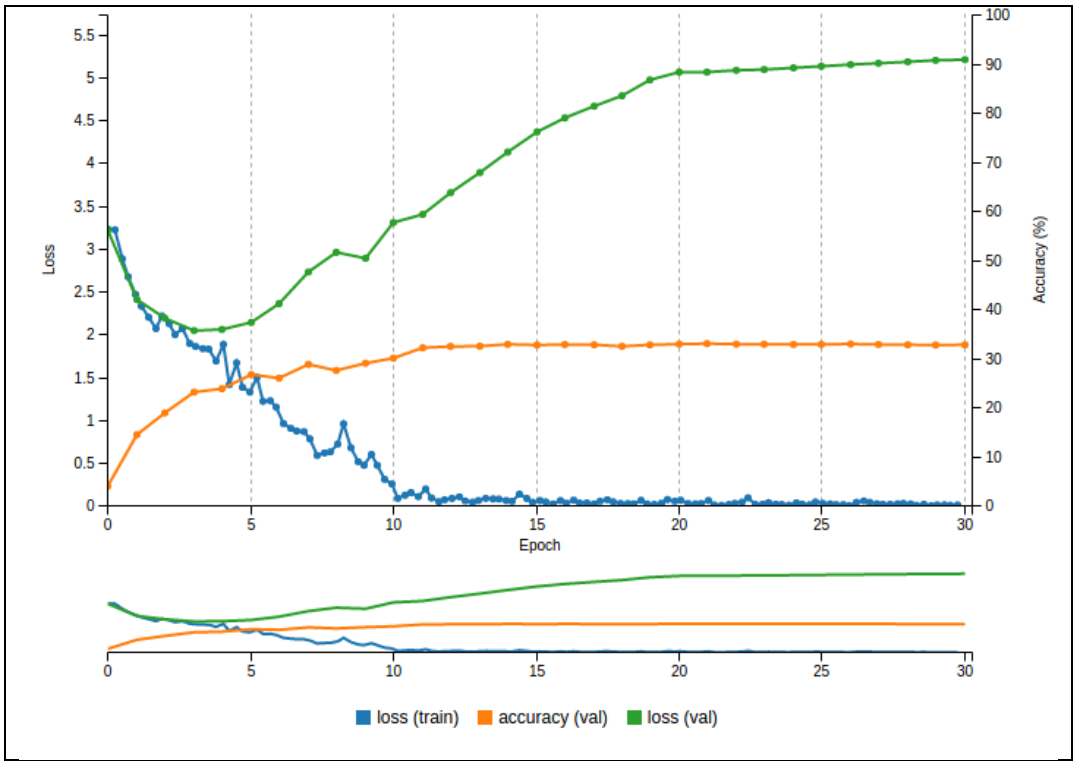
5.2. Performans Karşılaştırması

5.2.1. DIGITS Eğitim Grafikleri

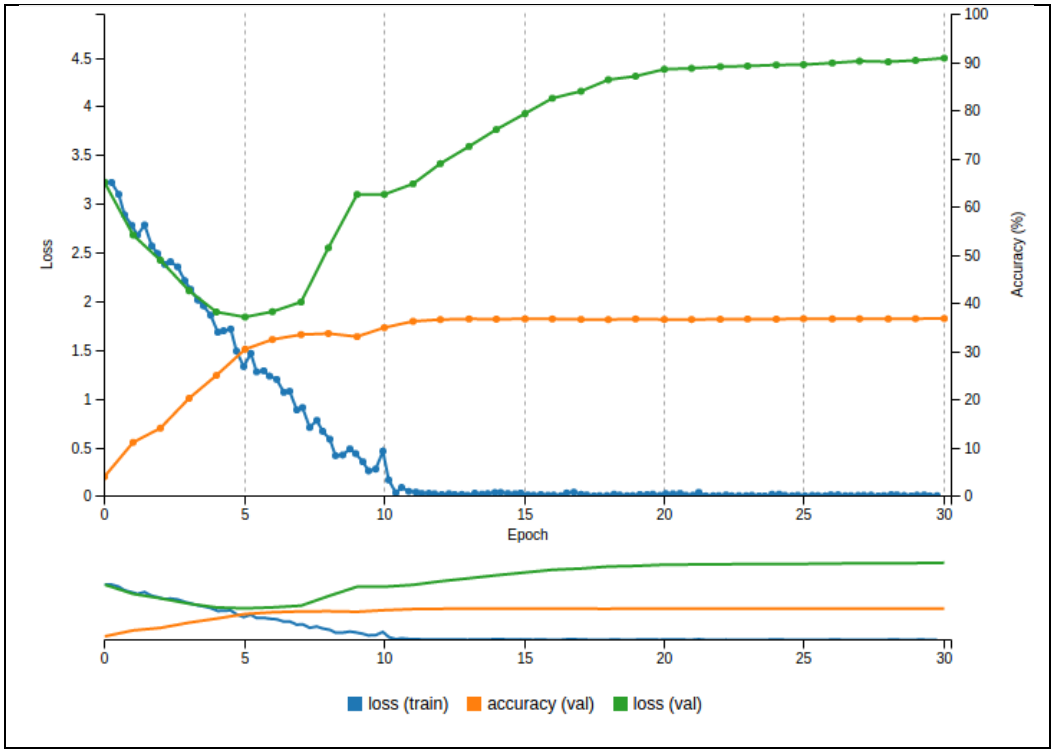
Sistemin eğitim ve test aşamalarında DIGITS [49] arayüzü kullanılmıştır. Bu arayüz, eğitim esnasında ve sonrasında, veri kümesinin model ile uyumluluğunu grafikler ile göstermektedir. Grafiklerde eğitim kaybı, doğrulama kaybı ve doğruluk oranı gerçekleştirilen devir sayısına göre gösterilmektedir. En yüksek doğruluk oranının çok ölçekli uzam-zamansal RGB modelde (Şekil 5.8), en düşük doğruluk oranını ise uzamsal modelde (Şekil 5.5) olduğu görülmektedir. Uzam-zamansal RGB (zamansal parçacıklar birleşimi) ve uzam-zamansal RGB (zamansal parçacıklar farkı) modelli ise ortalama bir performans göstermiştir (Şekil 5.6, Şekil 5.7).



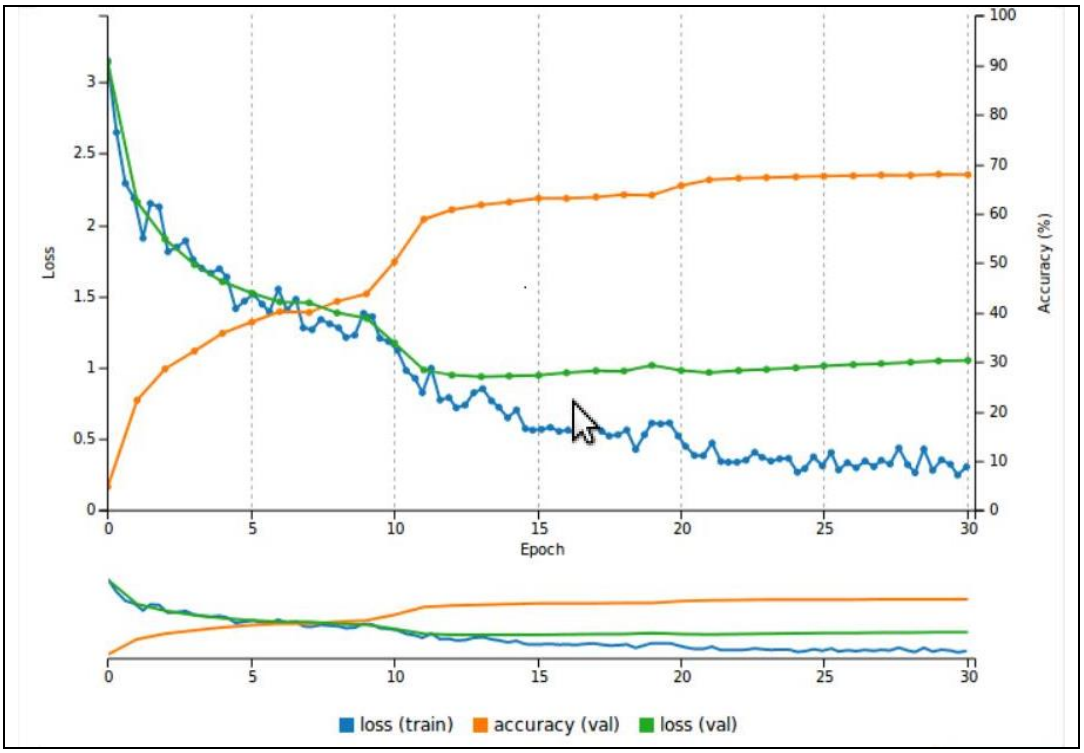
Şekil 5.5: Uzamsal model.



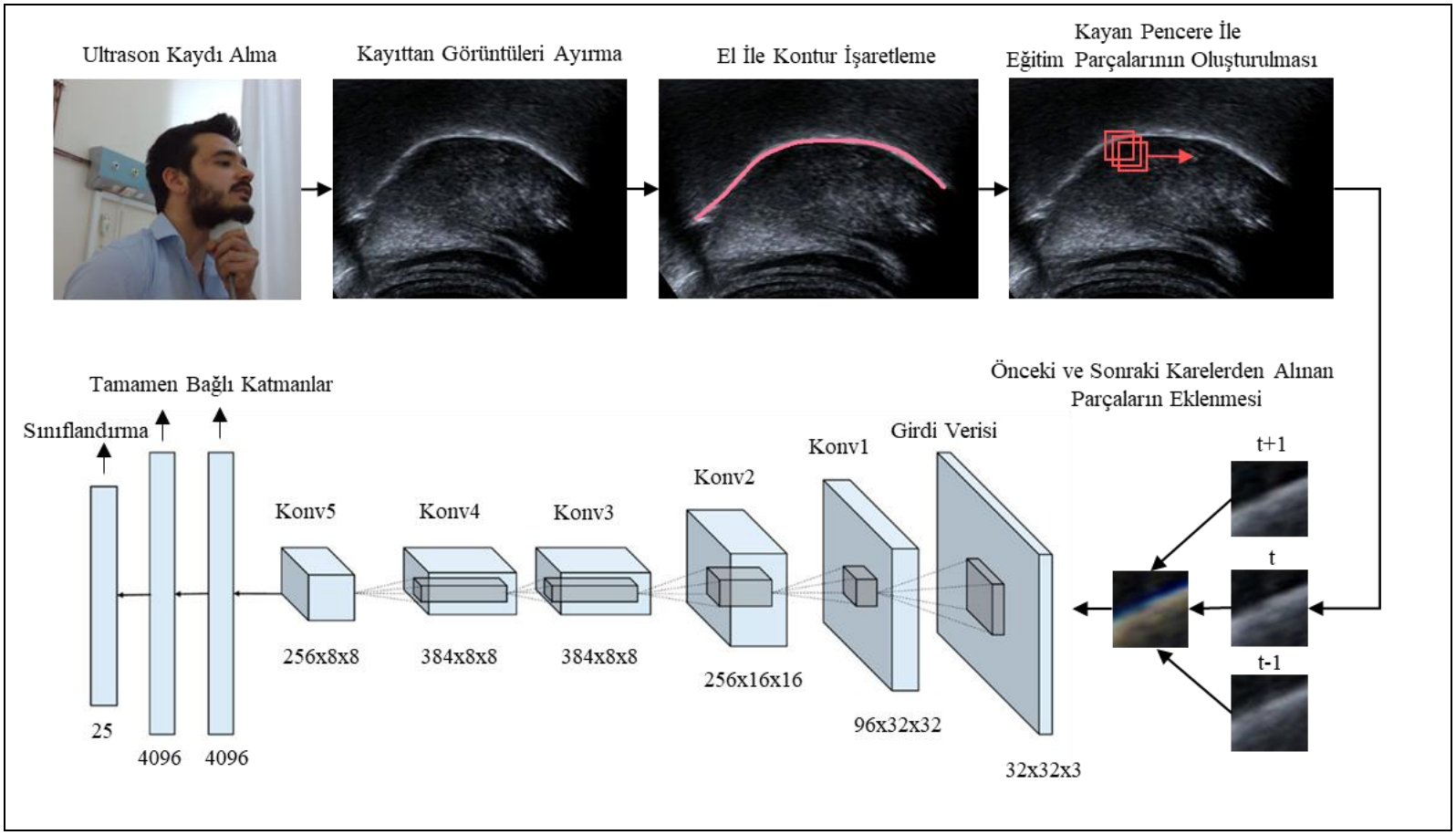
Şekil 5.6: Uzam-zamansal RGB (zamansal parçacıklar birleşimi) model.



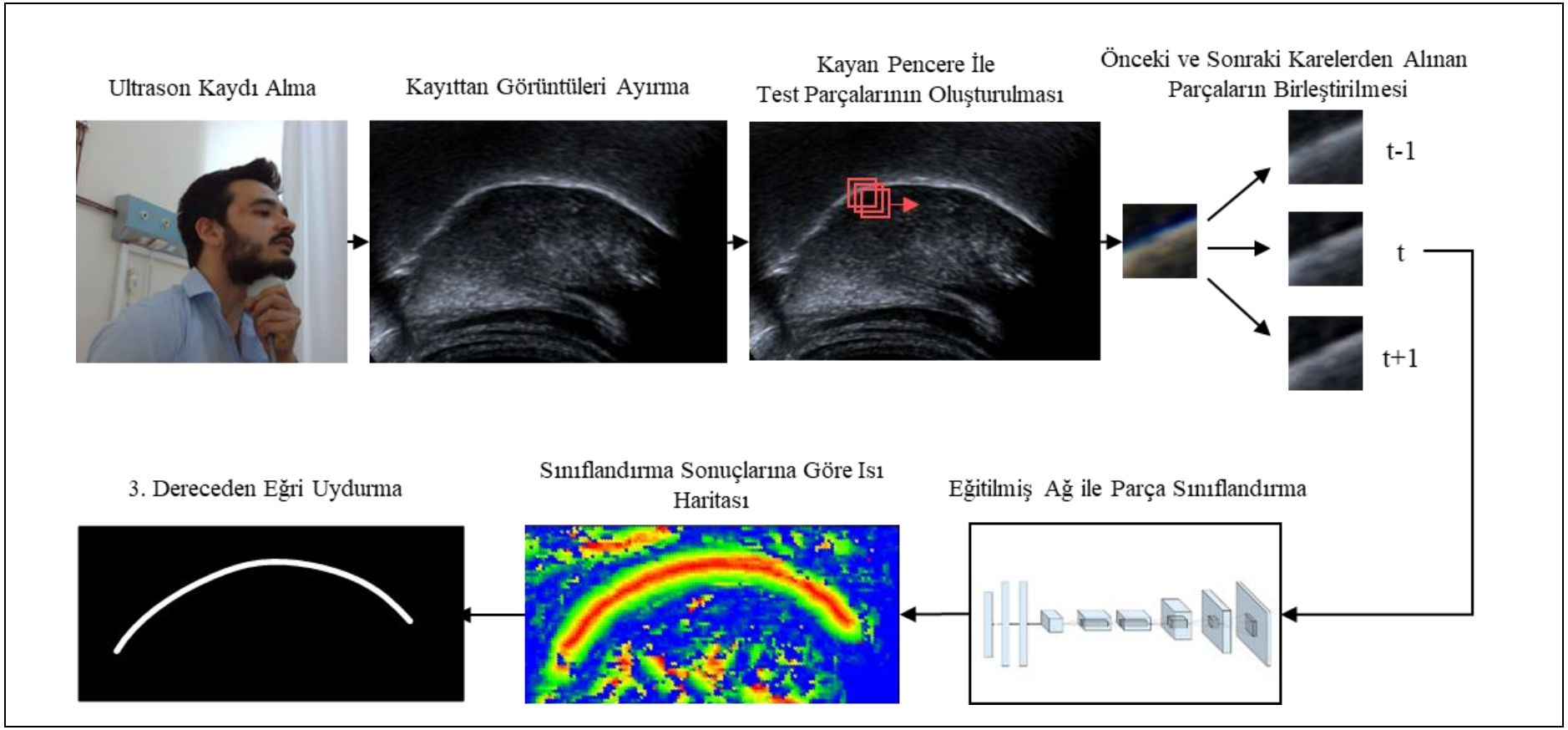
Şekil 5.7: Uzam-zamansal RGB (zamansal parçacıklar farkı) model



Şekil 5.8: Çok ölçekli uzam-zamansal RGB model.



Şekil 5.9: CNN ile dil konturu tespiti sisteminin eğitim aşamaları.



Şekil 5.10: CNN ile dil konturu tespiti sisteminin test aşamaları.

5.3. 3. Dereceden Eğri Uydurma

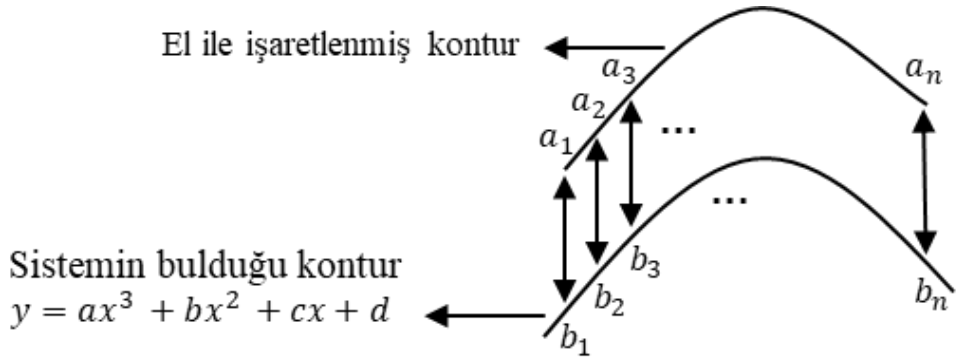
Dil konturu tespiti için CNN tabanlı uzam-zamansal, tek ölçekli-çok ölçekli olmak üzere birçok deneme yapılmıştır. Bu denemeler kontrollü olarak yapılan ölçümler ile birbirleri ile karşılaştırılmıştır. Son sistemde, RGB kanalları ile uzam-zamansal model kullanılmıştır.

$$y = ax^3 + bx^2 + cx + d \tag{5.2}$$

Sistem çıktı olarak ultrason görüntüsü üzerindeki her bir 32x32'lik kare için bir sınıf üretmektedir. Kontura en yakın üç sınıfa (0, 1, 2) denk gelen parçacıkların merkez noktaları kullanılarak regresyon [48] ile Denklem 5.2'deki gibi 3.dereceden bir polinom uydurulmuştur. Denkleme göre x ve y görüntünün satır ve sütunlarını, "a, b, c, d" ise denklemin sabit katsayılarını göstermektedir.

5.4. Deneyler ve Sonuçlar

Deneyler, Bölüm 4'te bahsedilen veri kümesinden alınan 927 görüntü ile gerçekleştirilmiştir. Bu görüntülerin 489'ü sistemin eğitimi için 438'i ise testi için kullanılmıştır. Eğitim ve test veri kümeleri farklı oturumlardan alınmıştır. Eğitimde 489 görüntüden alınan rastgele parçacıklar ile her bir sınıf için 10.000 parçacık ve 25 sınıf olmak üzere toplamda 25.000 parçacık kullanılmıştır. Düzenlenmiş AlexNet ağının eğitimi Nvidia'nın Tesla K20Xm [50] model ekran kartı ile gerçekleştirilmiştir. Eğitim ve testler Nvidia'nın DIGITS [49] arayüzü kullanılarak gerçekleştirilmiştir. Eğitim süresince veri kümesi ile 30 devir yapılmıştır ve eniyileme için SGD [40] kullanılmıştır.



Şekil 5.11: CNN ile dil konturu tespiti yapan sistemin ölçüm metriği.

Gerçekleştirilen sistemin başarımı Denklem 5.3 kullanılarak ölçülmüştür. Denkleme göre, a_0 noktası $0 \leq i \leq n$ olmak üzere manuel olarak işaretlenmiş dil konturunun x eksenindeki ilk değeri ve karşılığı olan y değeri ikilisinden oluşan noktayı, a_n ise aynı şekilde elde edilmiş konturun son noktasını göstermektedir. b_0-b_n arasında kalan noktalar ise gerçekleştirilen sistemin ürettiği dil konturunun üzerine Denklem 5.2'de gösterildiği gibi uydurulmuş polinomun, $x = 0$, $x = n$ değerleri arasında kalan noktalarını göstermektedir. Hata hesabı için her bir a, b noktaları arası uzaklık ölçülerek toplanmış ve kontur üzerindeki toplam nokta sayısına bölünmüştür (Denklem 5.3). Sistemde kullanılan verilerin elde edildiği ultrason cihazının parametreleri ile gerçekleştirilen hesaplamalar sonucu 1 pikselin 0.241 milimetreye denk geldiği görülmüştür. Tablo 5.1'de sistemin başarımı milimetre ve piksel cinsinden ölçülmüştür.

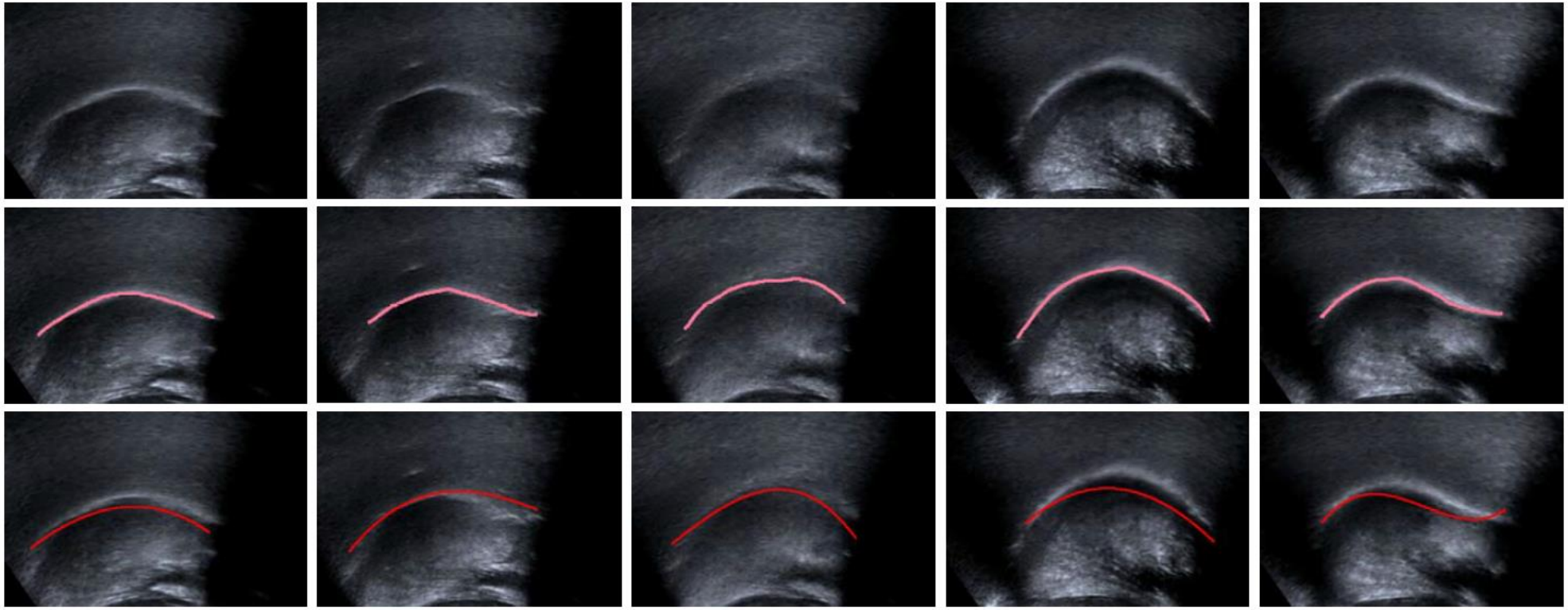
$$Ortalama\ Hata\ (px) = \frac{\sum_{i=0}^{n}(a_i - b_i)}{n} \tag{5.3}$$

Deneyler, sistemin ürettiği sonuçlardan en iyi ilk 1-7 arası sonuç kullanılarak sıralı şekilde yapılmıştır. Buna göre en düşük hatanın ilk 3 sınıf kullanıldığında çıktığı gözlenmiştir (Tablo 5.1). Sistemin eğitim ve test aşamasında kullanılan tüm görüntülerin çözünürlüğü 960x720'dir.

Tablo 5.1: CNN ile kontur bulan sistemin sonuçları.

İlk n Skor	Ortalama Hata (px)
1	21.42
2	20.97
3	20.79
4	20.92
5	21.78
6	21.99
7	22.29

5.5. Örnek Çıktılar



Şekil 5.12: CNN ile dil konturu tespiti sonuçları. 1.satırda ham ultrason görüntüleri, 2.satırda el ile işaretlenmiş dil konturları, 3.satırda ise sistemin bulduğu dil konturları gösterilmiştir.

6. LSTM İLE DİL KONTURU TESPİTİ

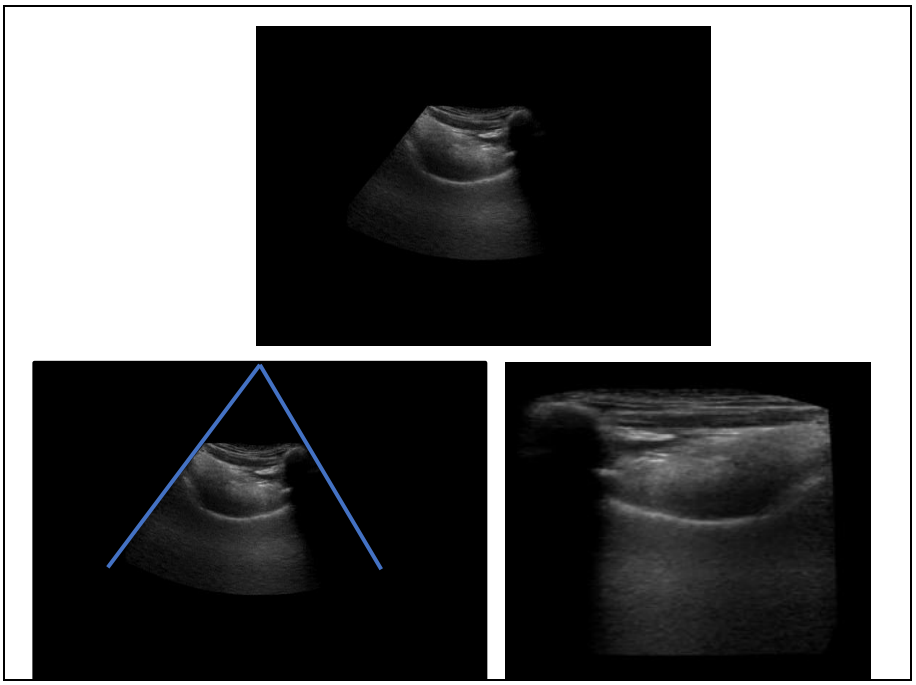
LSTM'ler zamana bağlı veri setleri için kullanılan RNN'lerin özelleşmiş bir hali olan sinir ağlarıdır. Çalışmada kullanılan ultrason görüntüleri, videolardan oluşmaktadır. Zamana bağlı olan bu görüntüler, bir bütün halinde anlam ifade etmektedir. Örnek olarak bir oturumda kaydedilen ultrason görüntülerinden oluşmuş video baştan sona değerlendirildiğinde anlamlı olacaktır. Videonun ardışık karelerindeki dil konturunun pozisyonu kareler arasında benzerlik göstermektedir. Evrişimsel yapay sinir ağları görüntü üzerindeki uzamsal özellikleri çıkarır. Görüntünün öncesinde geçen görüntüler ile ilgilenmez. LSTM ise içinde veri devamlılığını sağlayan döngüler barındırmaktadır. Bu döngüler sayesinde veri t anından $t+n$ anına taşınabilir. Bu döngülerin her biri sinir ağlarına benzer RNN hücreleri gibidir. T anındaki veri girdisi alınırken, önceki t anlarına ait bilgileri de girdi olarak alınmaktadır. Bu sayede zamana bağlı veriler bir bütün halinde değerlendirilebilmektedir. LSTM'ler bu özellikleri ile evrişimsel sinir ağlarından ayrışmaktadır.

Dil konturunu tespit eden derin öğrenme tabanlı yöntemler incelendiğinde tümünün ultrason dili kontur tespitini genel bir kontur tespit problemi olarak ele aldığı ve piksel sınıflandırma işi olarak düşündüğü görülmektedir. Oysa ki ultrason görüntüsü, yapısı ve oluşumu gereği diğer görüntülerden oldukça farklıdır, bu nedenle ultrason görüntülerinden dil konturu tespiti için özel yöntemler kullanılmalıdır. Ultrason görüntülerinde dil konturları uzamsal ve zamansal olarak düzenlidir, belirli bir düzene göre hareket ederler, ultrason konturu görüntünün her bir kolonunu bir kez keser. Tezin bu bölümünde, bu bilgiler kullanılarak geliştirilen sistemde iki tür LSTM ağı kullanılmıştır. Bunlar, uzamsal kontur izleyici ve uzam-zamansal nokta izleyici olarak isimlendirilmiştir [1]. Uzamsal kontur izleyici, ultrason görüntüsünün her karesi için 2B dil konturlarını pürüzsüz bir şekilde bulmaya çalışır. Uzam-zamansal nokta izleyici ise uzamsal izleyicinin çıktısını kullanır ve zaman içinde verilen dil konturunu izlemeye çalışır ve elde edilen konturların zamansal pürüzsüzlüğünü sağlar. Bu iki ağ, aynı anda beklenen uzamsal ve zamansal pürüzsüzlük kısıtlamalarını karşılamak için sırayla birbirlerinin çıktılarını da girdi olarak alarak eğitilir.

6.1. Kutupsal Koordinat Sistemine Dönüşüm

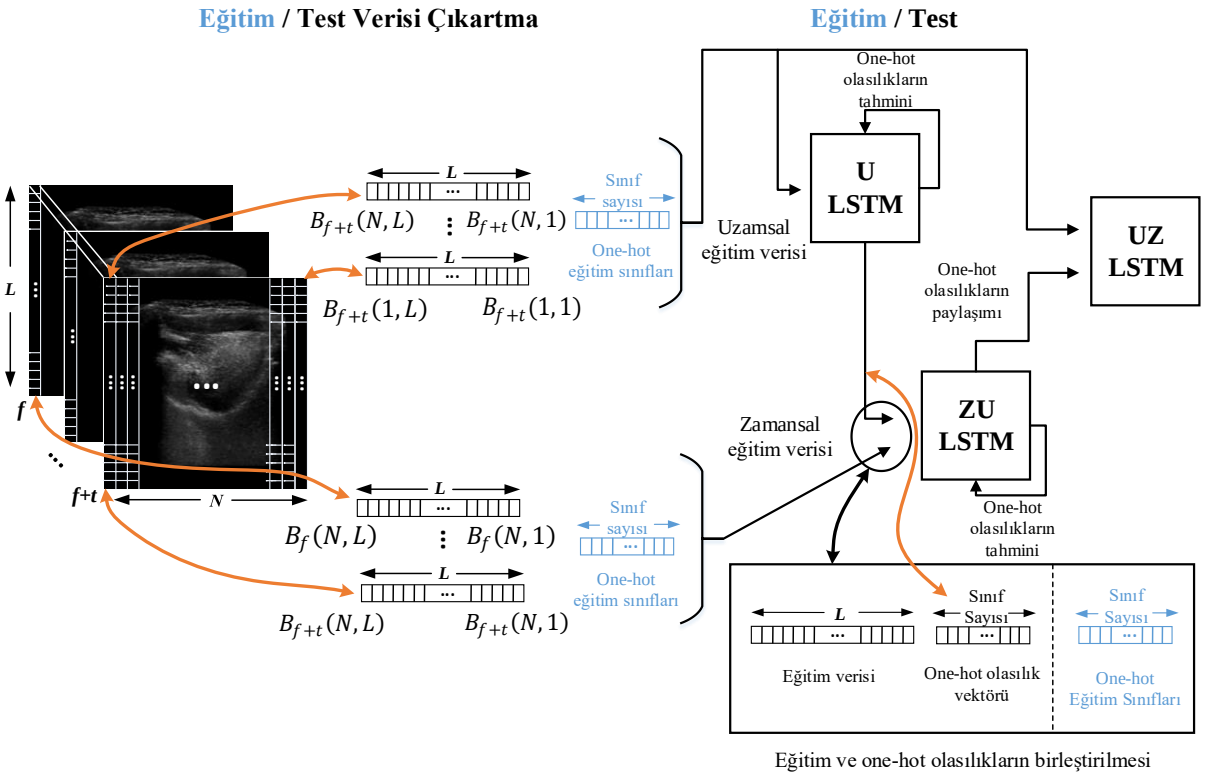
Bu bölümde uygulanan yöntemin temel motivasyonu, ultrason görüntüsünün oluşumu ve doğasından gelen avantajların görüntü işlemede kullanılma fikridir. Ultrason cihazından alınan görüntüler dikdörtgensel koordinat sistemindedir. Tez çalışmasında uygulanan sistemde görüntüler, kutupsal koordinat sistemine dönüştürülmüştür. Bu sayede ultrason cihazından çıkan her bir dalga görüntü matrisindeki bir kolona denk getirilmiş ve dil konturu, görüntünün her bir sütununu bir kez kesecek şekle gelmiştir. Eğitim ve test verileri sıralı bir şekilde LSTM ağına uygun hale getirilmiştir. Ağın her bir zaman adımına bir kolon yani bir ultrason dalgası denk gelmektedir.

Ultrason cihazından alınan görüntüler 960x720 çözünürlüğündedir. Şekildeki çizilmiş doğruların kesişim noktası kutupsal dönüşümün merkezi kabul edilerek dönüşüm gerçekleştirilerek 250x250 boyutunda kutupsal görüntü elde edilmiştir. Bu görüntünün her bir kolonu ultrason cihazından yansıyan bir dalgaya denk gelmektedir.



Şekil 6.1: LSTM ile dil konturu tespiti yapan sistemin eğitim ve test aşamaları.

6.2. Sistemin Eğitimi ve Testi



Şekil 6.2: LSTM ile dil konturu tespiti yapan sistemin eğitim ve test aşamaları.

Sistemde kullanılan eğitim verilerinin etiketleri one-hot vektör haline dönüştürülmüştür. Bu dönüşüm, Denklem 6.1 kullanılarak şu şekilde gerçekleştirilmiştir: D, vektörün uzunluğu, L ise x koordinatı yani derinlik olmak üzere tüm eğitim görüntüleri analiz edilerek dil konturunun geçtiği en küçük ve en büyük L koordinatları tespit edilmiştir. Bu koordinatların farkı alınarak one-hot vektörün uzunluğu hesaplanmıştır. Buna göre her bir kolon için konturun geçtiği satırın vektörde denk geldiği eleman 1 diğer tüm elemanlar 0 olmak üzere eğitim verisi etiketleri oluşturulmuştur.

$$D = \max(L) - \min(L) \tag{6.1}$$

Şekil 6.2'de görüldüğü gibi sistemin eğitimi için uzamsal ve zamansal olmak üzere iki tip eğitim verisi oluşturulmuştur. Uzamsal eğitim verileri her bir görüntü ayrı ayrı kullanılarak elde edilmiştir. Görüntülerin kolonlarının piksel değerleri eğitim verisi matrisinin satırlarına denk gelmektedir (Denklem 6.2). Zamansal eğitim verileri ise her bir kolon için tüm görüntülerin ilgili kolona denk gelen kolonları alınıp birleştirilerek oluşturulmuştur (Şekil 6.2).

Sistem 3 adet LSTM ağından oluşmaktadır. İlk ağ görüntülerin kolonları üzerinde çalışan ve konturun görüntüler içerisinde pürüzsüz bir şekilde tespit edilmesini sağlan uzamsal U-LSTM'dir. Bu ağ, Denklem 6.2'deki eğitim verileri ile eğitilir, burada her bir satır görüntü üzerindeki bir kolona yani ultrason yansıtıcısından çıkan bir dalgaya denk gelmektedir.

$$SF_f = \begin{bmatrix} B_f(1,1) & B_f(1,2) & \cdots & B_f(1,L) \\ B_f(2,1) & B_f(2,2) & & B_f(2,L) \\ \vdots & & \ddots & \vdots \\ B_f(N,1) & B_f(N,2) & \cdots & B_f(N,L) \end{bmatrix} \tag{6.2}$$

Sonrasında aynı veriler ile eğitilmiş model üzerinde tahmin yapılarak her bir satır için one-hot olasılık vektörü elde edilir. Bu vektör elemanları toplamı 1 olan, her bir eleman için 0-1 arasında bir olasılık tutan modelin çıktı vektörüdür. Sistemin ikinci ağı zaman-uzamsal ZU-LSTM'dir (Şekil 6.2). Ağ girdi olarak Denklem 6.3'te

gösterildiği gibi zamansal veri vektörünü ve U-LSTM'in one-hot olasılıklarını birleştirerek alır.

$$TSF_N \begin{bmatrix} B_f(N,1) & \cdots & B_f(N,L) & OH_f(N,1) & \cdots & OH_f(N,D) \\ B_{f+1}(N,1) & & B_{f+1}(N,L) & OH_{f+1}(N,1) & & OH_{f+1}(N,D) \\ \vdots & \vdots & & \vdots & & \\ B_{f+1}(N,1) & \cdots & B_{f+1}(N,L) & OH_{f+t}(N,1) & \cdots & OH_{f+t}(N,D) \end{bmatrix} \tag{6.3}$$

ZU-LSTM ağı eğitilerek model oluşturulur. Bu model üzerinde eğitim verileri ile tahmin yapılarak bir önceki aşamada olduğu gibi one-hot olasılık vektörü elde edilir. Sistemin son ağı uzam-zamansal UZ-LSTM ağdır. Bu ağ Denklem 6.3'te gösterildiği gibi girdi olarak uzamsal eğitim verisi matrisi ile ZU-LSTM ağının one-hot olasılık vektörünün birleşmiş halini alır. Denklemlerde B_f ultrason görüntüleri üzerindeki koordinatları ifade ederken OH_f one-hot vektörünü ifade etmektedir. Son ağ da Denklem 6.4'te gösterildiği gibi eğitilerek model oluşturulmuştur. Eğitim aşaması sonunda U-LSTM, ZU-LSTM ve UZ-LSTM aracılığı ile 3 adet model elde edilmiştir.

$$STF_f = \begin{bmatrix} B_f(1,1) & \cdots & B_f(1,L) & OH_f(1,1) & \cdots & OH_f(1,D) \\ B_f(2,1) & & B_f(2,L) & OH_f(2,1) & & OH_f(2,D) \\ \vdots & \vdots & & \vdots & & \\ B_f(N,1) & \cdots & B_f(N,L) & OH_f(N,1) & \cdots & OH_f(N,D) \end{bmatrix} \tag{6.4}$$

Sistemin test aşamasında eğitimde oluşturulmuş 3 adet model kullanılmıştır. Test görüntülerinden eğitim aşamasında olduğu gibi veriler çıkartılmıştır. Sırası ile önce U-LSTM ağının ürettiği model ile tahmin yapılarak one-hot çıktılar alınmış, ardından bu çıktılar zamansal test verileri ile birleştirilerek ZU-LSTM ağının ürettiği model üzerinde tahmin yapılmıştır. Buradan elde edilen one-hot olasılık vektörü ise uzamsal test verileri ile birleştirilerek sistemin son ağı olan UZ-LSM ağının ürettiği modele verilip son tahmin yapılarak dil konturu için noktalar tespit edilmiştir.

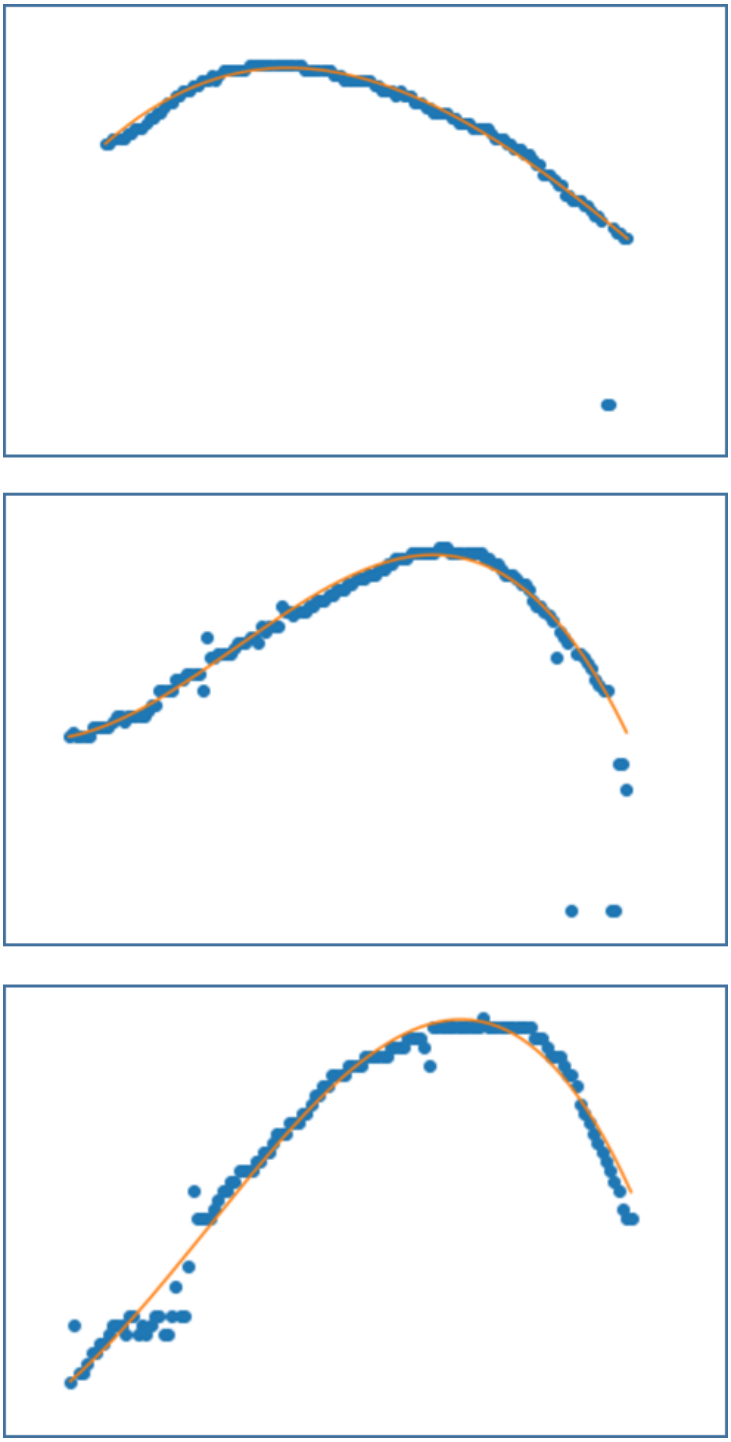
Sistemin ürettiği noktaların bazılarında kopmalar olduğu gözlenmiştir. Bu kopmaları engellemek için RANSAC [41] algoritması kullanılmıştır. Bir sonraki bölümde bu işlem detaylı olarak anlatılmıştır.

6.3. RANSAC İle Eğri Uydurma

RANSAC, matematiksel bir modelin parametrelerini, model üzerinde bir etkisi olmayan veya modeli bozan aykırı değerlerin tahmin edilmesi için geliştirilmiş bir algoritmadır. Birçok problem üzerinde uygulanarak hızlı ve oldukça başarılı sonuçlar elde edilmektedir. RANSAC için aykırı değer tespit etme algoritması denilir. Yineleme tabanlı bir algoritmadır. Her zaman en iyi sonucu üretmemektedir fakat yineleme sayısı ve parametreler iyi ayarlandığında olabildiğinde en iyiye yakın sonucu üretmektedir. Algoritma ilk olarak 1981 yılında Fischler ve Bolles tarafından SRI International'da [41] yayınlanmıştır. Bu yayında RANSAC algoritması yer belirleme problemi üzerinde kullanılmıştır.

Algoritmaya sistemin modeli verilir. Girdi olarak verilen örneklerde, modele uymayan örnekler elenerek tüm girdiler ile algoritma yeniden çalıştırılır. Yinelemeler ile algoritma çalıştırılarak aykırı girdiler elenir.

Tez çalışmasında, LSTM ile dil konturu bulma bölümünde geliştirilen sistemde, son aşamada görüntünün her bir kolonu için kontur olabilecek nokta tahmin edilmektedir. Normalde görüntünün ardışık kolonlarındaki kontur pozisyonu çok fazla yer değiştirmemelidir. Sistemin ürettiği çıktılar incelendiğinde bazı kolonlar arasındaki kontur noktası pozisyonları arasında olması gerekenden fazla bir fark olduğu gözlemlenmiştir. Bu sebepten dolayı kontur üzerinde olarak bulunan noktaların aykırı olanlarını tespit etmek ve elemek için RANSAC algoritması kullanılmıştır. Bulunan noktalar algoritmaya verilerek dil konturu için 3.dereceden bir polinom uydurulmuştur. Böylece sistemin başarımı oldukça yüksek seviyelere çıkartılmıştır. Şekil 6.3’te sistem tarafından bulunan noktalar RANSAC algoritmasından geçirildikten sonra elde edilen 3. dereceden örnek eğriler gösterilmiştir. Görüldüğü gibi eğri, kopan noktalardan etkilenmemiş kontur doğru bir şekilde tespit edilmiştir.



Şekil 6.3: LSTM'in bulduğu noktalar ve bu noktalara RANSAC ile uydurulmuş eğriler gösterilmiştir.

6.4. Ağ Parametreleri ve Uzam-zamansal Ağ Adedi

Tez çalışmasında geliştirilen sistemin birçok parametresi vardır. Derin ağlar eğitilirken seçilen gizli katman sayısı, zaman adımı sayısı, kayıp fonksiyonu, eniyileyici türü, toptan boyutu gibi parametreler sistemin başarımını doğrudan etkilemektedir. Bu parametrelerin başarım üzerinde etkileri, girdi olarak verilen verinin türüne, miktarına göre değişim göstermektedir. Bu sebepten dolayı her sistem kendi içerisinde değerlendirilmelidir. Tez çalışmasında gerçekleştirilen sistemin parametreleri de farklı değerler ile çok sayıda eğitim ve test yapılarak en iyi sonucu veren parametre kombinasyonu seçilmiştir.

Sistemin başarımını etkileyen bir diğer faktör ise uzamsal ve zamansal ağların sayısıdır. Bölüm 5.2'de bahsedildiği gibi sistem, uzamsal ve zamansal ağların çıktılarının paylaşılarak sıralı bir şekilde eğitilmesi ile oluşturulmuştur. Dolayısı ile kullanılan ağ sayısı da yapılan deneyler sonucunda en iyi sonucu üretecek şekilde seçilmiştir.

6.4.1. Ağ Parametreleri Seçimi

Sistem, LSTM ağları kullanılarak oluşturulmuştur. Eğitim aşamasında, zaman adımı sayısı, gizli katman sayısı ve toptan boyutu parametreleri değiştirilerek eğitim ve testler gerçekleştirilmiştir. Bu testler, uzamsal bir LSTM ağı ile gerçekleştirilmiştir. Sonrasında ise seçilen parametreler tüm ağların eğitimi için kullanılmıştır.

Tablo 6.1: Ağ parametreleri ve hata oranları.

Zaman Adımı	Gizli Katman	Toptan	MSD(mm)
7	75	20	1.62
7	100	20	3.81
7	25	20	2.20
7	75	70	3.86
7	75	110	4.32
10	75	20	4.44
12	75	20	4.67
13	75	20	5.69

Tablo 6.1'de gösterildiği gibi, ilk olarak zaman adımı ve toptan sayısı sabit tutularak gizli katman sayısı değiştirilip en uygun olan seçilmiştir. Sonrasında gizli katman sayısı ve zaman adımı sayısı sabit tutularak toptan sayısı değiştirilip en uygun olan seçilmiştir. Son olarak da toptan sayısı ve gizli katman sayısı belirlenen değerlerde sabit tutularak zaman adımı sayısı değiştirilmiş ve onun da en uygun değeri seçilmiştir. Zaman adımı sayısı 7, gizli katman sayısı 75 ve toptan sayısı da 20 olarak belirlenmiştir. Belirlenen bu değerler sistemde kullanılan tüm ağlarda aynı verilmiştir.

6.4.2. Uzam-zamansal Ağ Adedi Seçimi

Sistem, uzamsal ve zamansal ağların birbirlerinin çıktıları ile eğitimi üzerine kurulmuştur. Dolayısı ile kaç adet sıralı ağ kullanılacağı başarımı doğrudan etkilemektedir. Bu sayıyı en uygun şekilde seçmek için bir dizi eğitim ve testler gerçekleştirilmiştir.

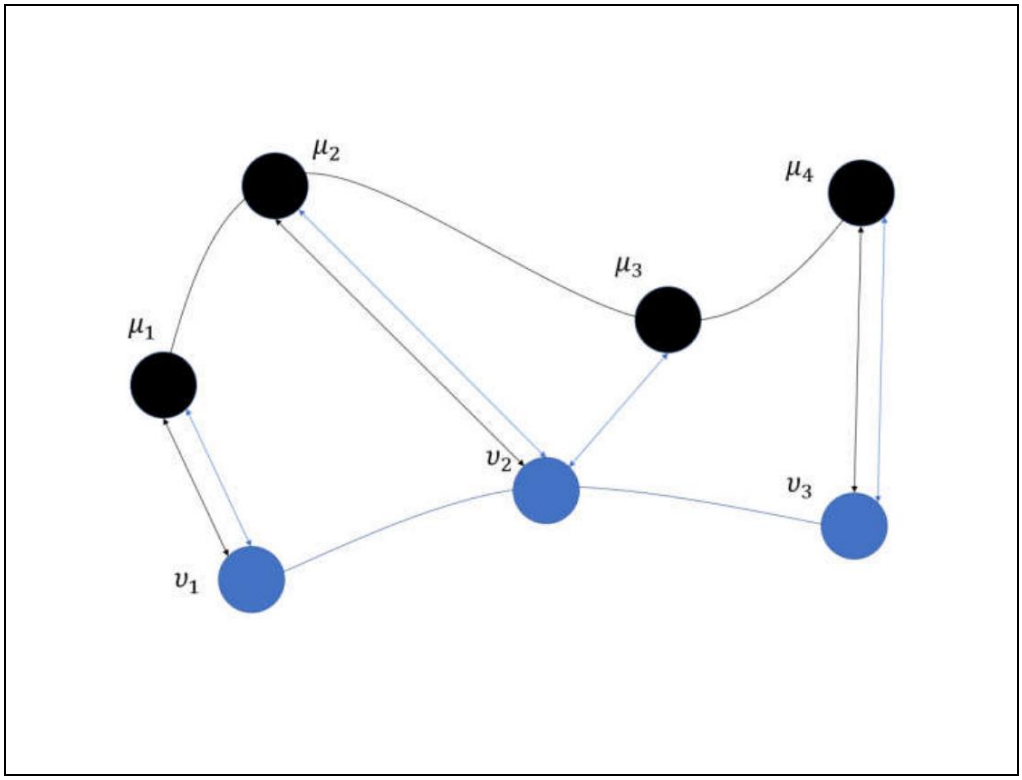
Tablo 6.2: Uzamsal ve zamansal ağ sayısı ve hata oranları.

Uzam-zamansal	Zaman-uzamsal	MSD (mm)
1	1	0.7
2	2	0.63
3	3	0.61
4	4	0.58
5	5	0.62

Tablo 6.2'de farklı sayılarda ağl kullanılarak testler gerçekleştirilmiştir. Sistem en düşük hata oranını 4 uzam-zamansal, 4 zaman-uzamsal ağ olmak üzere toplamda 8 ağ kullanıldığı durumda elde etmiştir (0.58 mm). 1 uzam-zamansal ve 1 zaman-uzamsal ağ kullanıldığı durumda ise hata oranı 0.7 mm'dir. Buradaki seçim 1-1 toplamda 2 ağ kullanılacak şekilde yapılmıştır. Sebebi ise çok fazla ağ kullanımının getireceği eğitim ve test maliyetine rağmen başarımda ciddi oranda bir iyileşme olmamasıdır.

6.5. Deneyler ve Sonuçlar

Deneyler, Bölüm 4'te bahsedilen veri kümesinden alınmış görüntüler ile gerçekleştirilmiştir. 972 görüntünün 544'ü sistemin eğitimi için, 428'i ise testi için kullanılmıştır. Eğitim ve test aşamalarında kullanılan görüntüler, farklı oturumlardan alınmıştır. Sistem geliştirilirken Keras [42] kütüphanesi, TensorFlow [43] arka ucu ile birlikte kullanılmıştır. Kullanılan tüm LSTM ağlarının gizli katman [44] sayısı 75, kayıp fonksiyonu Categorical Crossentropy [47], en iyileyicisi Adam [29] ve toptan boyutu [45] 20'dir. Tüm veri kümesi üzerinde 30 devir [46] yapılmıştır.



Şekil 6.4: İki kontur üzerindeki nokta sayısı eşit olmasa bile MSD metriği ile ölçüm gerçekleştirilebilir.

Sistemin başarımının ölçümü için literatürdeki birçok çalışmada olduğu gibi MSD ölçüm metriği kullanılmıştır. Bu metrik Denklem 6.5'te gösterilmiştir. Denklemde görüldüğü gibi U ve V konturları arasındaki ortalama uzaklık piksel cinsinden ölçülür. Bu ölçüm iki konturun da her x değeri için bir y değeri olmasına ihtiyaç duymaz, eksik değerler olsa bile ölçüm yapabilir (Şekil 6.4). Buna göre, her iki konturun üzerindeki noktalar, sırası ile diğer konturda karşılık geldiği en yakın noktadan çıkartılarak elde edilen fark, tüm noktalar için toplanır. Bu işlem U-V ve V-U olarak gerçekleştirilir.

$$MSD(U,V) = \frac{1}{m+n}\left(\sum_{i=1}^{m} \min_j |v_i - u_j| + \sum_{i=1}^{m} \min_j |u_i - v_j|\right) \quad (6.1)$$

Elde edilen toplam değer, iki kontur üzerindeki tüm noktaların sayısına bölünerek ölçüm gerçekleştirilmiş olur.

Tez çalışmasının başarımının ölçümü içi 4 çeşit deney yapılmıştır. Bunlar: Sistemde kullanılan ağların birbiri üzerindeki etkilerini ölçen deney, sistem sonuçlarını literatürdeki yöntemler ile karşılaştıran deney, 5 katlamalı çapraz doğrulama ve referans noktası deneyleridir.

6.5.1. Tez Çalışmasındaki Ağların Birbiri Üzerlerindeki Etkileri

Sistem 3 adet LSTM ağından oluşmaktadır. Bu ağlar zamansal ve uzamsal bilgiler barındırmakta ve birbirlerinin çıktıları ile beslenmektedirler. İlk deney, dil konturu tespit etmede uzamsal ve zamansa bilgilerin önemini anlamak ve bu bilgilerin sonuçlar üzerinde etkisini gözlemlemek için gerçekleştirilmiştir.

Tablo 6.3: LSTM ile dil konturu bulan sistemde kullanılan ağların birbirlerinin başarımları üzerindeki etkileri.

Ağ	MSD (mm)	MSD (px)	İyileşme (%)
Z-LSTM	0.96	1.45	-
U-LSTM	0.42	0.63	-
ZU-LSTM	0.39	0.59	59.3
UZ-LSTM	0.34	0.52	17.46

Tablo 6.3'te gösterilen Z-LSTM zamansal LSTM'dir. Bu ağ sistemde kullanılmamıştır, fakat zamansal bilginin ne kadar etkisi olduğunu gözlemlemek için bu ağ da deneylerin içerisine dahil edilmiştir. Sonuçlarda Z-LSTM in tek başına 0.96 mm, U-LSTM'in ise 0.42 mm hata ile çalıştığı görülmektedir. U-LSTM'nin one-hot bilgileri ZU-LSTM'ye verildiğinde ağın 0.39 mm hata yaptığı görülmüştür, bu da U-

LSTM'nin ZU-LSTM sonuçlarını %59 iyileştirdiği anlamına gelmektedir. ZU-LSTM one-hot olasılık sonuçları ise UZ-LSTM'ye verildiğinde UZ-LSTM'de %17.46'lık bir iyileşme olduğu gözlemlenmiştir. Bu sonuçlar, sistemin bulduğu noktalara RANSAC algoritması uygulanmadan elde edilmiştir.

6.5.2. Literatürdeki Yöntemler ile Karşılaştırma

İkinci deney sistemin sonuçlarını literatürdeki en başarılı yöntemler ile karşılaştırmak amacı ile gerçekleştirilmiştir. Literatürdeki derin öğrenme tabanlı en başarılı yöntemler [21] ve [7]'dir. Aktif kontur tabanlı çalışmalardan ise [4] en başarılı yöntemdir. Tez çalışmasında 1 piksel 0.68 milimetreye denk gelirken [21]'de 0.295 milimetreye, [7] 'de 0.638 milimetreye ve [4]'te 0.295 milimetreye denk gelmektedir. Tablo 6.4'te görüldüğü gibi, tez çalışması 0.28 milimetre hata ile çalışırken, [7] 0.91 milimetre, [21] 1.0 milimetre, [4] ise 0.5 milimetre hata ile çalışmaktadır.

Tablo 6.4: LSTM ile dil konturu bulma sonuçlarının literatürdeki yöntemlerin sonuçları ile karşılaştırılması.

Yöntem	MSD(mm)
Tez Çalışması	0.28
SegNet & sU-net [13]	0.91
DBN [21]	1.0
Aktif Kontur [4]	0.5

6.5.3. 5 Katlamalı Çapraz Doğrulama

Sistemin eğitim ve testinde kullanılan veri kümeleri her ne kadar iki oturumdan oluşmuş olsa da kullanılan denekler aynı olduğu için LSTM'in denekleri öğrenme ihtimali bulunmaktadır. Bunu gözlemleyebilmek için 5 adet veri kümesi oluşturulmuştur. Her bir veri kümesinde bir denek eğitimde kullanılmayıp yalnızca test aşamasında kullanılmıştır. 5 denek için 5 veri kümesi oluşturularak 5 katlamalı çapraz doğrulama deneyleri gerçekleştirilmiştir. Deney sonuçları Tablo 6.5'te gösterilmiştir. Başarımda Bölüm 6.5.1'deki deneye göre ciddi bir düşüş olmaması,

tasarlanan ağın ultrason görüntülerinden dil konturu tespiti için uygun olduğunu göstermektedir.

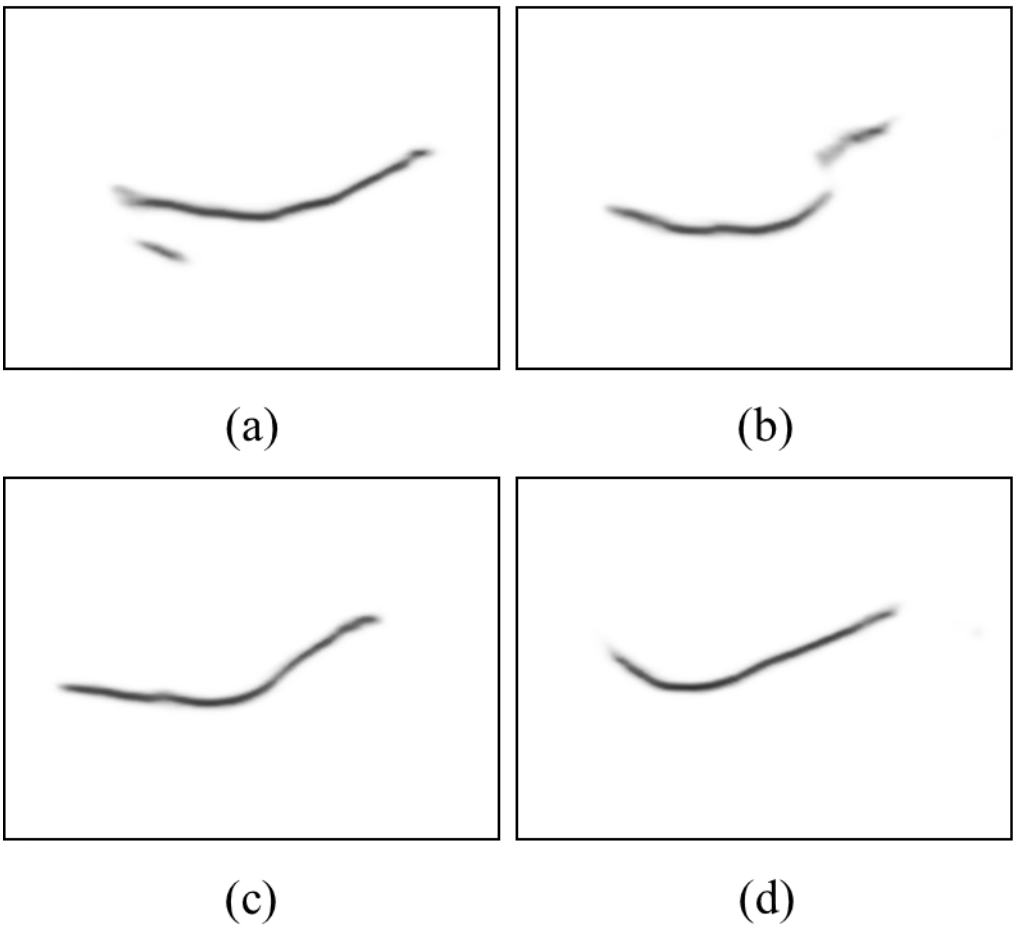
Tablo 6.5: 5 katlamalı çapraz doğrulama deneyleri sonuçları.

Ağ	1 MSD(mm)	2 MSD(mm)	3 MSD(mm)	4 MSD(mm)	5 MSD(mm)
Z-LSTM	1.55	1.43	1.70	1.62	1.50
U-LSTM	0.68	0.76	0.65	0.89	0.83
ZU-LSTM	0.63	0.72	0.59	0.77	0.79
UZ-LSTM	0.59	0.57	0.52	0.72	0.62

6.5.4. Referans Noktası Deneyi

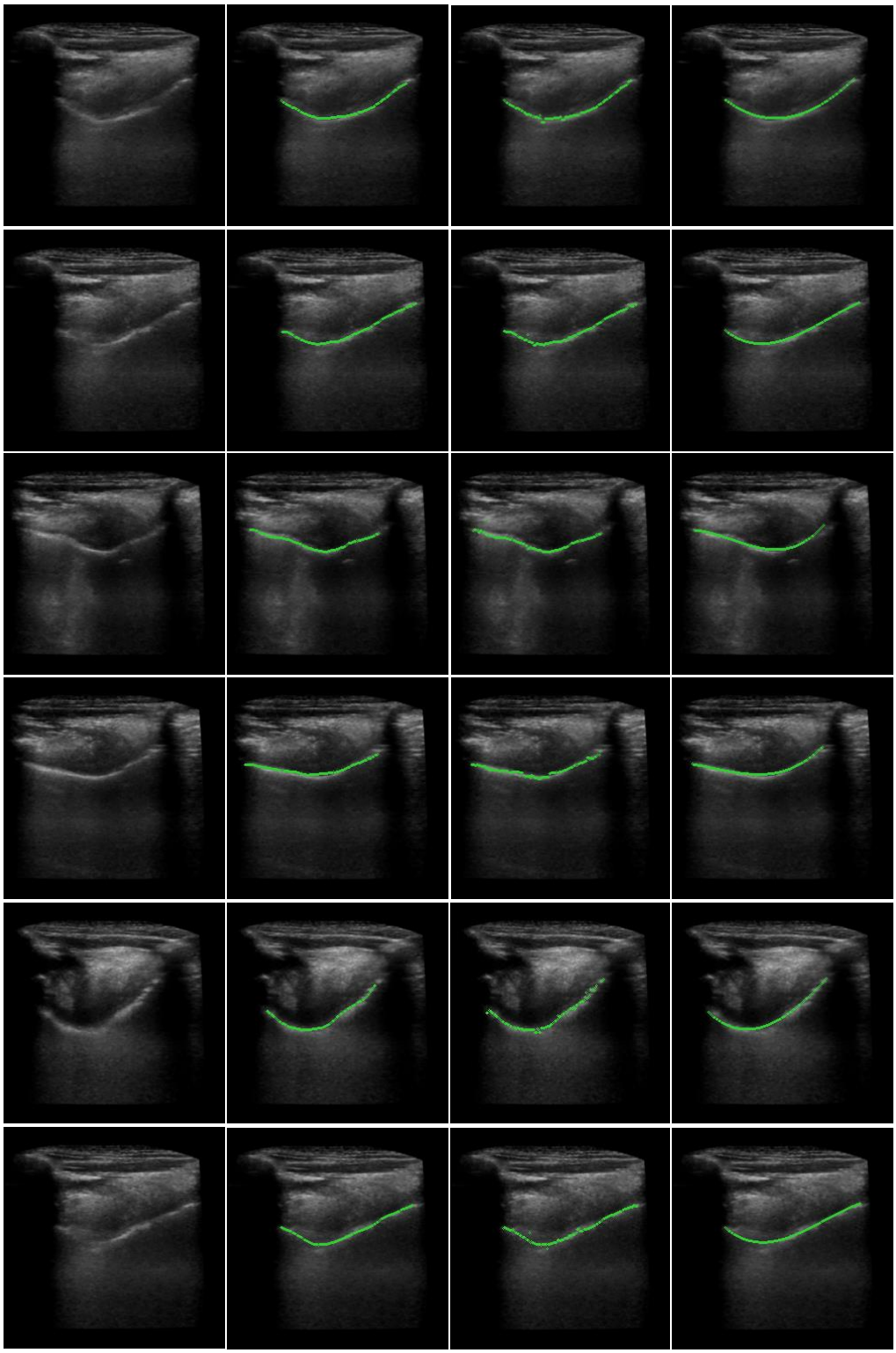
Literatürde, biyomedikal görüntüler üzerinde iyi çalıştığı bilinen U-Net ağı, tez çalışmasında kullanılan veri kümesi ile çalıştırılmıştır. Şekil 6.5'te örnek sistem çıktıları gösterilmiştir. U-Net ağı hızlı bir şekilde çıktı üretmesine karşın piksel sınıflandırması yapmaktadır. Şekil 6.5'te (a) görüntüsünde görüldüğü gibi dil konturunun altında da kontur gibi ufak bir bölge bulunmuştur, dil konturu hiçbir zaman bu şekilde bölünemez, aslında bu durum U-Net gibi piksel sınıflandırması yapan sistemlerin, ultrason görüntülerinden dil konturu tespiti için kullanılmasının uygun olmadığını göstermektedir. Tez çalışmasında gerçekleştirilen LSTM tabanlı yöntem, her bir ultrason görüntüsü matrisindeki kolonda bir adet kontur noktası olabilir veya hiç olamaz varsayımı ile (a) görüntüsündeki soruna çözüm getirmiştir.

U-Net ağı, Bölüm 6.5'te bahsedilen veri kümesinin aynısı ile eğitilip testleri gerçekleştirilmiştir. Denklem 6.5'te gösterilen formüle göre hata hesabı gerçekleştirilmiştir. Sistemin hata oranı 0.99 mm olarak ölçülmüştür. Tez çalışmasında gerçekleştirilen LSTM tabanlı sistemin gerisinde kalmıştır. Tez çalışmasındaki sistemin RANSAC uygulanmadan elde ettiği başarım Tablo 6.3'te gösterilmiştir. (0.34 mm)



Şekil 6.5: Örnek U-Net çıktıları.

6.6. Örnek Çıktılar



Şekil 6.6: LSTM ile dil konturu tespit etme sonuçları, 1.sütun ham görüntüyü, 2.sütun el ile işaretlenmiş konturu, 3.sütun sistemin bulduğu noktaları 4.sütun noktalara RANSAC ile uydurulan 3.derece polinomu göstermektedir.

7. SONUÇLAR

Ultrason görüntüleri, ultrason teknolojisinin doğası gereği, görüntünün her yerinde, izlenmek istenen organa ait eşit bilgi içermemektedir. Dolayısı ile görüntü işlenirken bu bilgi göz önünde bulundurulmalıdır. Bunun yanı sıra oldukça fazla gürültü içeren görüntülerdir, bu durum da ultrason görüntüleri üzerinden kontur tespit etme işlemini oldukça zor bir hale getirmektedir. Literatürdeki dil konturu tespiti yapan çalışmalar incelendiğinde birçoğunun konturu tespit etmek için sistemi manuel olarak ilklendirdiği, bir dizi müdahaleler ile kontur tespiti işlemi gerçekleştirdiği görülmektedir. Manuel müdahale yapmayan sistemlerin ise problemi klasik bir piksel sınıflandırma olarak ele aldığı görülmektedir. Bu tez çalışmasında ultrason teknolojisinin doğasına uygun, yalnızca ultrason görüntülerinden kontur tespiti yapan LSTM tabanlı bir sistem geliştirilmiştir. Bu sistemden bağımsız olarak CNN kullanarak parçacık tabanlı dil konturu tespiti yapan ikinci bir sistem daha geliştirilmiştir. Literatürdeki yöntemlere kıyasla oldukça başarılı sonuçlar elde edilmiştir. LSTM ile geliştirilen çalışmanın motivasyonunu ultrason görüntülerinin elde ediliş şekli oluşturmuştur. Buna göre ultrason görüntüleri dikdörtgensel koordinat sisteminden kutupsal koordinat sistemine çevrilerek görüntünün her bir kolonu ultrason yansıtıcısından çıkan bir dalgaya denk gelecek şekle getirilmiştir, böylece görüntünün her alanı dil konturuna dair eşit seviyede bilgi içerir hale getirilmiştir. Geliştirilen sistem ultrason görüntüleri üzerinde çalışması için özelleşmiş bir sistemdir. Bu alanda yapılmış çalışmalara bakıldığında hepsinin problemi bir piksel sınıflandırması veya sıradan bir kontur tespit etme problemi olarak ele aldığı görülmektedir. Çalışmayı literatürdeki yöntemlerden ayıran bir diğer önemli farkı ise uzamsal ve zamansal bilgileri bir arada kullanmasıdır. Ultrason görüntülerinin video karelerinden elde edilmesinden dolayı, kareler arasında kontur pozisyonunun çok fazla değişmeyeceği göz önünde bulundurularak uzamsal ve zamansal bilgiler birleştirilmiş ve bu şekilde hata oranı düşük bir kontur tespiti gerçekleştirilmiştir. Tez çalışması bu özellikleri ile literatürdeki diğer çalışmalardan ayrılmaktadır.

Çalışmada dil konturu tespiti için geliştirilmiş ikinci sistem, CNN kullanarak parçacık tabanlı dil konturu tespiti gerçekleştirmektedir. Ultrason görüntüsünden elde edilen parçacıklara kontura uzaklığına göre sınıf atanıp nesne tanımada başarılı olarak çalıştığı bilinen AlexNet ağı eğitilmiş ve oluşan model kullanılarak kontur tespiti

gerçekleştirilmiştir. Bu çalışmada da uzamsal ve zamansal bilgiler bir arada kullanılmıştır.

Sistem sonuçları literatürdeki ultrason görüntülerinden dil konturu tespiti yapan aktif kontur ve derin öğrenme tabanlı çalışmalar ile karşılaştırılmıştır. Geliştirilen sistemin şimdiye kadar yapılmış tüm çalışmalardan daha iyi sonuçlar vermiştir. Bunun yanında hız performansı olarak da gerçek zamana yakın bir performans göstermektedir. Sistem manuel ilklendirme veya müdahale gerektirmeksizin tam otomatik bir şekilde kontur tespiti yapmaktadır. Gelecek çalışmalar için sistemin hız performansı daha da arttırılarak gerçek zamanlı hale getirilmesi ve konuşma terapisi için gerçekleştirilecek bir uygulamada kullanılması planlanmaktadır.

KAYNAKLAR

[1] Aslan E., & Akgul Y. S., (2019), "Tongue Contour Tracking in Ultrasound Images With Spatiotemporal LSTM Networks", In German Conference on Pattern Recognition.

[2] Abadi M., Barham P., Chen J., Chen Z., Davis A., Dean J., ... & Kudlur M. (2016), "Tensorflow: A system for large-scale machine learning", In 12th Symposium on Operating Systems Design and Implementation (16), 265-283.

[3] Akgul Y. S., Kambhamettu C., & Stone M. (1999), "Automatic extraction and tracking of the tongue contours", IEEE Transactions on Medical Imaging, 18(10), 1035-1045.

[4] Li M., Kambhamettu C., & Stone M. (2005), "Automatic contour tracking in ultrasound images", Clinical linguistics & phonetics, 19(6-7), 545-554.

[5] Stone M. (2005), "A guide to analysing tongue motion from ultrasound images", Clinical linguistics & phonetics, 19(6-7), 455-501.

[6] Xu K., Yang Y., Stone M., Jaumard-Hakoun A., Leboullenger C., Dreyfus G., ... & Denby B. (2016), "Robust contour tracking in ultrasound tongue image sequences", Clinical linguistics & phonetics, 30(3-5), 313-327.

[7] Wen S. (2018), "Automatic Tongue Contour Segmentation using Deep Learning", Doctoral dissertation, Université d'Ottawa/University of Ottawa.

[8] Lai K. F., & Chin R. T. (1995), "Deformable contours: Modeling and extraction", IEEE Transactions on pattern analysis and machine intelligence, 17(11), 1084-1090.

[9] Aslan E., Dumlu N., & Akgül Y. S. (2018), "Tongue contour extraction from ultrasound images using image parts", In 2018 26th Signal Processing and Communications Applications Conference (SIU), 1-4, IEEE.

[10] Hinton G. E. (2012), "A practical guide to training restricted Boltzmann machines", In Neural networks: Tricks of the trade, 599-619.

[11] Hoskins P. R., Martin, K., & Thrush, A., (2010), "Diagnostic ultrasound: physics and equipment. Cambridge University Press".

[12] Gérard J. M., Perrier P., & Payan Y. (2006), "3D biomechanical tongue modeling to study speech production".

[13] Mozaffari M. H., & Wen S., & Wang N., & Lee WS., (2019), "Real-time

Automatic Tongue Contour Tracking in Ultrasound Video for Guided Pronunciation Training", Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, 302-309.

[14] Chum O., Matas J., & Kittler J. (2003), "Locally optimized RANSAC", In Joint Pattern Recognition Symposium, 236-243, Springer.

[15] Akgul Y. S., Kambhamettu C., & Stone M. (2000), "A task-specific contour tracker for ultrasound", In Proceedings IEEE Workshop on Mathematical Methods in Biomedical Image Analysis. 135-142, IEEE.

[16] Amini A. A., Weymouth T. E., & Jain R. C. (1990), "Using dynamic programming for solving variational problems in vision", IEEE Transactions on Pattern Analysis & Machine Intelligence, (9), 855-867.

[17] A. Blake., & M. Isard., (1998), "Active Contours", Springer-Verlag.

[18] Cohen L. D., (1991), "On active contour models and balloons", Computer Vision, Graphics, Image Processing-Image Understanding, 53(2):211–218.

[19] Geiger D., Gupta A., Costa L. A., & Vlontzos J. (1995), "Dynamic programming for detecting, tracking, and matching deformable contours", IEEE Transactions on Pattern Analysis & Machine Intelligence, (3), 294-302.

[20] Gunn, S. R. & Nixon M. S. (1997), "A robust snake implementation; a dual active contour", IEEE Transactions on pattern analysis and machine intelligence, 19(1), 63-68.

[21] Fasel I., & Berry J. (2010), "Deep belief networks for real-time extraction of tongue contours from ultrasound during speech", In 2010 20th International Conference on Pattern Recognition (1493-1496), IEEE.

[22] Noble A., & Boukerroui D., (2006), “Ultrasound image segmentation: a survey,” IEEE Trans. Med. Imaging, 25(8), 987–1010.

[23] Abel J., Allen B., Burton S., Kazama M., Kim B., Noguchi M., ... & Gick B. (2015), "Ultrasound-enhanced multimodal approaches to pronunciation teaching and learning", 43(3), Canadian Acoustics.

[24] Bernhardt M. B., Bacsfalvi P., Adler-Bock M., Shimizu R., Cheney A., Giesbrecht N., ... & Radanov B. (2008), "Ultrasound as visual feedback in speech habilitation: Exploring consultative use in rural British Columbia, Canada", Clinical Linguistics & Phonetics, 22(2), 149-162.

[25] Preston J. L., McCabe P., Rivera-Campos A., Whittle J. L., Landry E., & Maas E. (2014), "Ultrasound visual feedback treatment and practice variability for

residual speech sound errors", Journal of Speech, Language, and Hearing Research, 57(6), 2102-2115.

[26] Burckhardt C. B., (1978), "Speckle in ultrasound B-mode scans", IEEE Transactions on Sonics and ultrasonics, 25(1), 1-6.

[27] Bridal S. L., Correas J. M., Saied, A. M. E. N. A., & Laugier P. (2003), "Milestones on the road to higher resolution, quantitative, and functional ultrasonic imaging", Proceedings of the IEEE, 91(10), 1543-1561.

[28] Drozdzal M., Vorontsov E., Chartrand G., Kadoury S., & Pal C. (2016), "The importance of skip connections in biomedical image segmentation", In Deep Learning and Data Labeling for Medical Applications (179-187), Springer, Cham.

[29] Lawson E., Stuart-Smith J., Scobbie J. M., Nakai S., Beavan D., Edmonds F., ... & Esling J. (2015), "Seeing Speech: an articulatory web resource for the study of phonetics", website.

[30] De Boor C., De Boor, C., Mathématicien E. U. (1978), "A practical guide to splines", (27, 325), New York: springer-verlag.

[31] Krizhevsky A., Sutskever, I., & Hinton G. E., (2012), "Imagenet classification with deep convolutional neural networks", In Advances in neural information processing systems (1097-1105).

[32] Hochreiter S., & Schmidhuber J. (1997), "Long short-term memory", Neural computation, 9(8), 1735-1780.

[33] Myers R. H., (1990), "Classical and modern regression with applications", (2), Belmont, CA: Duxbury press.

[34] Cha S. H., (2007), "Comprehensive survey on distance/similarity measures between probability density functions", City, 1(2), 1.

[35] Kass M., Witkin A., & Terzopoulos D., (1988), "Snakes: Active contour models. International journal of computer vision", 1(4), 321-331.

[36] Duda R. O., Hart P. E., & Stork D. G. (2012), "Pattern classification", John Wiley & Sons.

[37] Heijmans H. J., (1994), "Morphological image operators", (4), Boston: Academic Press.

[38] Badrinarayanan V., Kendall A., & Cipolla R., (2017), "Segnet: A deep convolutional encoder-decoder architecture for image segmentation", IEEE

transactions on pattern analysis and machine intelligence, 39(12), 2481-2495.

[39] G. Castellano L. Bonilha L. M. Li, & F. Cendes., (2004), "Texture analysis of medical images", Clin. Radiol.

[40] Bottou L., (2010), "Large-scale machine learning with stochastic gradient descent", In Proceedings of COMPSTAT'2010 (177-186), Physica-Verlag HD.

[41] Fischler M. A., & Bolles R. C., (1981), "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography", Communications of the ACM, 24(6), 381-395.

[42] F. Chollet., (2015), "Keras: Deep Learning library for Theano and TensorFlow," GitHub Repos.

[43] M. Abadi., (2016), "TensorFlow: A System for Large-Scale Machine Learning TensorFlow: A system for large-scale machine learning," in 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI '16).

[44] Panchal G., Ganatra A., Kosta Y. P., & Panchal, D., (2011), "Behaviour analysis of multilayer perceptrons with multiple hidden neurons and hidden layers", International Journal of Computer Theory and Engineering, 3(2), 332-337.

[45] Schmeiser B., (1982), "Batch size effects in the analysis of simulation output. Operations Research", 30(3), 556-568.

[46] Baumjohann W., Paschmann G., Nagai T., & Lühr H., (1991), "Superposed epoch analysis of the substorm plasma sheet", Journal of Geophysical Research: Space Physics, 96(A7), 11605-11608.

[47] Zhang Z., & Sabuncu M. (2018), "Generalized cross entropy loss for training deep neural networks with noisy labels", In Advances in Neural Information Processing Systems (8778-8788).

[48] Kutner Michael H., (2005), "Applied linear statistical models", (5), Boston: McGraw-Hill Irwin.

[49] Yeager L., (2015), "DIGITS: the deep learning GPU training system." ICML 2015 AutoML Workshop.

[50] Lindholm E., (2008), "NVIDIA Tesla: A unified graphics and computing architecture", IEEE micro 28(2) 39-55.

[51] Richmond K., & Renals S., (2012), "Ultrax: An animated midsagittal vocal tract display for speech therapy", In Thirteenth Annual Conference of the International Speech Communication Association.

[52] Wilson I., & Gick B., (2006), “Ultrasound technology and second language acquisition research”, In Proceedings of the 8th Generative Approaches to Second Language Acquisition Conference (GASLA 2006) (148-152). Somerville, MA: Cascadilla Proceedings Project.

[53] Potočnik B., & Zazula D., (2002), “Automated analysis of a sequence of ovarian ultrasound images”, Part I: segmentation of single 2D images. Image and Vision Computing, 20(3), 217-225.

[54] Mora J., Lee, W. S., & Comeau G., (2007), “3D visual feedback in learning of piano posture”, In International Conference on Technologies for E-Learning and Digital Entertainment (763-771). Springer, Berlin, Heidelberg.

[55] Mora J., Lee W. S., Comeau G., Shirmohammadi S., & El Saddik A. (2006), “Assisted piano pedagogy through 3d visualization of piano playing”, In 2006 IEEE International Workshop on Haptic Audio Visual Environments and their Applications (HAVE 2006) (157-160). IEEE.

[56] Drowatzky J. N., (1981), “Motor learning: Principles and practices”, (No. BF295. D76.). Minneapolis, MN:: Burgess Publishing Company.

[57] Xu K., Gábor Csapó T., Roussel P., & Denby B. (2016), “A comparative study on the contour tracking algorithms in ultrasound tongue images with automatic re-initialization”, The Journal of the Acoustical Society of America, 139(5), EL154-EL160.

[58] Rastadmehr O., Bressmann T., Smyth R., & Irish J. C. (2008), “Increased midsagittal tongue velocity as indication of articulatory compensation in patients with lateral partial glossectomies”, Head & neck, 30(6), 718-726.

[59] Mozaffari M. H., & Lee W. S. (2017), “Freehand 3-D Ultrasound Imaging: A Systematic Review”, Ultrasound Med. Biol., 43(10), 2099–2124.

[60] Laporte C., & Ménard L. (2015), “Robust tongue tracking in ultrasound images: a multi-hypothesis approach”, In Sixteenth Annual Conference of the International Speech Communication Association.

[61] Tang L., Hamarneh G., & Bressmann T. (2011), “A machine learning approach to tongue motion analysis in 2d ultrasound image sequences”, In International Workshop on Machine Learning in Medical Imaging (151-158), Springer, Berlin, Heidelberg.

[62] Tang L., Bressmann T., & Hamarneh G. (2012), “Tongue contour tracking in dynamic ultrasound via higher-order MRFs and efficient fusion moves”, Medical image analysis, 16(8), 1503-1520.

[63] Fabre D., Hueber T., Bocquelet F., & Badin P. (2015), “Tongue tracking in

ultrasound images using eigentongue decomposition and artificial neural networks".

[64] McDicken W. N., & Anderson T. (2011), "Basic physics of medical ultrasound", Thrid Edit., (1), Elsevier Ltd.

[65] Saini K., Dewal M. L., & Rohit M. (2010), "Ultrasound imaging and image segmentation in the area of ultrasound: a review", International Journal of advanced science and technology, 24.

[66] Martin K. (2011), "Basic equipment, components and image production", In Clinical ultrasound (16-30). Churchill Livingstone.

[67] Shankar P. M. (2000), "A general statistical model for ultrasonic backscattering from tissues", IEEE transactions on ultrasonics, ferroelectrics, and frequency control, 47(3), 727-736.

[68] Mignotte M., & Meunier J. (2001), "A multiscale optimization approach for the dynamic contour-based boundary detection issue", Computerized Medical Imaging and Graphics, 25(3), 265-275.

[69] Mignotte M., Meunier, J., & Tardif J. C. (2001), "Endocardial boundary e timation and tracking in echocardiographic images using deformable template and markov random fields", Pattern Analysis & Applications, *4*(4), 256-271.

[70] Racine R., Walker G. A., Nadeau D., Doyon R., & Marois C. (1999), Speckle noise and the detection of faint companions, Publications of the Astronomical Society of the Pacific, 111(759), 587.

[71] Tao Z., Jaffe C. C., & Tagare H. D. (2003), "Tunnelling descent: A new algorithm for active contour segmentation of ultrasound images", In Biennial International Conference on Information Processing in Medical Imaging (246-257). Springer, Berlin, Heidelberg.

[72] Cardinal M. H. R., Meunier J., Soulez G., Thérasse É., & Cloutier G. (2003), "Intravascular ultrasound image segmentation: A fast-marching method", In International Conference on Medical Image Computing and Computer-Assisted Intervention (432-439). Springer, Berlin, Heidelberg.

[73] Zahalka A., & Fenster A. (2001), "An automated segmentation method for three-dimensional carotid ultrasound images", Physics in Medicine & Biology, 46(4), 1321.

[74] Paragios N., Jolly M. P., Taron M., & Ramaraj R. (2005), "Active shape models and segmentation of the left ventricle in echocardiography", In International Conference on Scale-Space Theories in Computer Vision (131-142). Springer, Berlin, Heidelberg.

[75] Xu K., Yang Y., Leboullenger C., Roussel P., & Denby B. (2016), “Contour-based 3D tongue motion visualization using ultrasound image sequences”, In 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (5380-5384). IEEE.

[76] Yilmaz A., Javed, O., & Shah M. (2006), “Object tracking: A survey”, Acm computing surveys (CSUR), 38(4), 13.

[77] Lowe D. G. (1999), “Object recognition from local scale-invariant features”, In iccv, 99(2), 1150-1157.

[78] Marr D., & Hildreth E. (1980), “Theory of edge detection”, Proceedings of the Royal Society of London. Series B. Biological Sciences, 207(1167), 187-217.

ÖZGEÇMİŞ

Enes ASLAN 1992 yılında Kocaeli'nin İzmit ilçesinde dünyaya geldi. İlk, orta ve lise öğrenimini Kocaeli'de tamamladı. 2010 yılında başladığı Gebze Teknik Üniversitesi Bilgisayar Mühendisliği Bölümü'ndeki lisans eğitiminden 2015 yılında mezun oldu. Ardından 2016 yılında Gebze Teknik Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Bölümü'nde yüksek lisans eğitimine başladı. İş hayatına ise 2015 yılında Kuveyt Türk Katılım Bankası'nda Yazılım Mühendisi olarak başladı ve Kıdemli Yazılım Mühendisi olarak halen devam etmektedir.

EKLER

Ek A: Tez Çalışması Kapsamında Yapılan Yayınlar

Aslan E., & Akgul Y. S. (2019), “Tongue Contour Tracking in Ultrasound Images With Spatiotemporal LSTM Networks”, In German Conference on Pattern Recognition.

Aslan E., Dumlu N., & Akgül Y. S. (2018), “Tongue contour extraction from ultrasound images using image parts”, In 2018 26th Signal Processing and Communications Applications Conference (SIU) (1-4). IEEE.