



T.C.
SELÇUK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ



GÖRÜNTÜ İŞLEME YÖNTEMLERİ İLE
YÜZEY ÜZERİNE OYULMUŞ
KARAKTERLERİN TANINMASI

Mahmut Sami YASAK

YÜKSEK LİSANS TEZİ

Mekatronik Mühendisliği Anabilim Dalını

Ağustos-2019
KONYA
Her Hakkı Saklıdır

TEZ KABUL VE ONAYI

Mahmut Sami YASAK tarafından hazırlanan “Görüntü İşleme Yöntemleri ile Yüzey Üzerine Oyulmuş Karakterlerin Tanınması” adlı tez çalışması 22/8/19 tarihinde aşağıdaki jüri üyeleri tarafından oy birliği / ~~oy çokluğu~~ ile Selçuk Üniversitesi Fen Bilimleri Enstitüsü Mekatronik Mühendisliği Anabilim Dalı’nda YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

Jüri Üyeleri

Başkan

Doç. Dr. Hasan Erdinç KOÇER

Danışman

Doç. Dr. Hasan Erdinç KOÇER

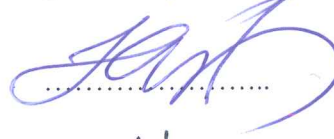
Üye

Dr. Öğr. Üyesi Kemal TÜTÜNCÜ

Üye

Dr. Öğr. Üyesi Onur İNAN

İmza



Yukarıdaki sonucu onaylarım.

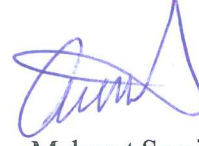
Prof. Dr. Mustafa YILMAZ
FBE Müdürü

TEZ BİLDİRİMİ

Bu tezdeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edildiğini ve tez yazım kurallarına uygun olarak hazırlanan bu çalışmada bana ait olmayan her türlü ifade ve bilginin kaynağına eksiksiz atıf yapıldığını bildiririm.

DECLARATION PAGE

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all materials and results that are not original to this work.



Mahmut Sami YASAK

Tarih: 22/08/2019

ÖZET

YÜKSEK LİSANS TEZİ

GÖRÜNTÜ İŞLEME YÖNTEMLERİ İLE YÜZEY ÜZERİNE OYULMUŞ KARAKTERLERİN TANINMASI

Mahmut Sami YASAK

**Selçuk Üniversitesi Fen Bilimleri Enstitüsü
Mekatronik Mühendisliği Anabilim Dalı**

Danışman: Doç. Dr. Hasan Erdinç Koçer

2019, 47 Sayfa

Jüri

**Doç. Dr. Hasan Erdinç KOÇER
Dr. Öğr. Üyesi Kemal TÜTÜNCÜ
Dr. Öğr. Üyesi Onur İNAN**

Günümüzde nesne tespit etme, nesne tanımlama, nesne sınıflandırma ve nesne takip etme problemleri için görüntü işleme algoritmaları ve bilgisayarlı görme içerisinde yapay zeka algoritmaları sıklıkla kullanılmaktadır. Nesnelerin tanınması probleminde çözme hızı ve başarı hassasiyet faktörleri dikkate alınarak yeni yöntemler sürekli olarak geliştirilmektedir.

İmalat sektöründe preslenerek şekil verilen metal parçalar üzerinde, tanımlanabilmeleri için marka, model ve seri numarası gibi bir çok amacı bulunan karakterler (harf, numara) ve semboller yer almaktadır. Bu karakterler ve semboller, buldukları yüzey üzerine oyulmuş veya kabartılmış şekillerde bulunmaktadır. Tez kapsamında metal yüzeyler üzerine oyulmuş karakterlerin sınıflandırılması için çalışma gerçekleştirilmiştir. Bu kapsamda, dairesel küçük çaplı metal cisimlerin yüzeyleri üzerinde oyulmuş halde bulunan karakterlerin sınıflandırılması için 2 farklı yaklaşımda bulunulmuştur.

Öncelikle gürültülerin giderilmesi amacıyla gri seviyedeki ham görüntülere Laplacian of Gaussian (LoG) filtresi uygulanmıştır. Ardından dairesel biçimdeki metal cisimlerin daire sınır çizgilerini belirlemek amacıyla Canny kenar tespit algoritması uygulanmış ve elde edilen kenar haritasına Dairesel Hough Dönüşümü (DHD) uygulanarak en dış, orta ve iç daire sınır çizgileri çizilmiştir. İç daire sınır çizgisinin içinde kalan bölge ön karakter bölgesi olarak belirlenmiştir.

İlk yaklaşımda, belirlenen ön karakter bölgesine Daugman'ın Rubber Sheet (DRS) modeli uygulanarak polar kartezyen dönüşümü yapılmıştır. Daha sonra dikdörtgen haldeki karakter bölgesi resmine Gabor filtresi uygulanmış ve sonucunda faz bilgisi elde edilmiştir. Faz bilgisine Ortalama Mutlak Sapma tekniği uygulanarak karakter bölgesine ait özellik vektörü elde edilmiştir. Son olarak ikili (binary) formata dönüştürülerek XOR operatörü ile şablon eşleştirme yapılmıştır.

İkinci yaklaşımda ise, yine belirlenen ön karakter bölgesine Maksimum Kararlı Ekstrimal Bölgeler (MSER) ve Kontür Genişliği Dönüşümü (SWT) algoritmaları uygulanmış ve karakter segmentasyonu gerçekleştirilmiştir. Segmente edilen karakter görsellerine yapay zekanın derin öğrenme yaklaşımlarından olan Evrişimli Sinir Ağları (ESA) kullanılarak tasarlanan model uygulanarak karakterler tanınmıştır. Karakter tanınması sağlanan dairesel metal cisimlerin aynı zamanda nesne sınıflandırma problemi çözülmüştür. Her iki yaklaşımda elde edilen başarı değerleri, doğruluk yüzdeleri ve işlem süreleri karşılaştırmalı olarak tezin sonuçlar bölümünde verilmiştir.

Anahtar Kelimeler: Evrişimli Sinir Ağları, Görüntü İşleme, Kontür Genişliği Dönüşümü, Maksimum Kararlı Ekstrimal Bölgeler, Nesne Sınıflandırma, Optik Karakter Tanıma, XOR Şablon Eşleştirme

ABSTRACT

MS THESIS

THE RECOGNITION OF ENGRAVING CHARACTER ON THE SURFACE WITH IMAGE PROCESSING METHODS

Mahmut Sami YASAK

**THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCE
OF SELÇUK UNIVERSITY
THE DEGREE OF MASTER OF SCIENCE
IN MECHATRONICS ENGINEERING**

Advisor: Assoc. Prof. Dr. Hasan Erdinç Koçer

2019, 47 Pages

Jury

Assoc. Prof. Dr. Hasan Erdinç KOÇER

Asst. Prof. Kemal TÜTÜNCÜ

Asst. Prof. Onur İNAN

Nowadays, object detection, object recognition, object classification and object tracking problems solve with image processing algorithms and commonly used in computer vision which is artificial intelligence algorithms. New methods are constantly being developed by considering the speed of solving problems and success sensitivity factors in the recognition of objects.

The sample metal products that shaped by press machine in manufacturing sector, these products includes characters (letter, number) or symbols that can be classified as brand, model and serial number. These characters are found in engraved or embossed forms on their metal surface. In this thesis, work was done for engraving characters on metal surfaces. In this context, two different approaches have been made for the classification of engraved characters on the surfaces of circular small diameter metal objects.

Firstly, the Laplacian of Gaussian (LoG) filter was applied to the gray level raw images in order to eliminate the noise. Then, Canny edge detection algorithm was applied in order to sharpen the circle boundary lines of metal objects, and then, outermost, middle and inner circle boundary lines were drawn by applying Circular Hough Transform (CHT) to the obtained edge map. The region inside the inner circle boundary line is determined as the pre-character region.

In the first approach, the polar cartesian transformation was performed by applying Daugman's Rubber Sheet (DRS) model to the specified pre-character region. Then, Gabor Filter (GF) was applied to rectangular character region image and phase information was obtained as a result. The characteristic vector of the character region was obtained by applying the Mean Absolute Deviation technique to the phase information. Finally, the binary format has been converted for pattern matching with the XOR operator.

In the second approach, Maximally Stable Extremal Regions (MSER) and Contour Width Transformation (SWT) algorithms were applied to the determined pre-character region and character segmentation was performed. The characters were recognized by applying the model designed by using Convolutional Neural Networks (CNN), one of the deep learning approaches of artificial intelligence, to segmented character images. At the same time, the classification problem of circular metal objects with character recognition is solved. Obtained performance metrics of both approaches that include accuracy rate and average processing times are given in the results section of the thesis.

Anahtar Kelimeler: Convolutional Neural Network, Image Processing, Maximally Stable Extremal Region, Stroke Width Transform, Object Classification, Optical Character Recognition, XOR Pattern Matching

ÖNSÖZ

Bu tez çalışmasında, nesne karakter tanıma veya nesne sınıflandırma problemlerin çözüm aşamaları içerisinde yer alan, nesnelerin ön karakter bölgelerinin bulunması ve şablon eşleştirmenin gerçekleştirilmesi, ön karakter bölgeler içerisindeki karakterlerin görüntü içerisindeki segmentasyonu ve karakterlerin tanınması için görüntü işleme algoritmaları uygulanmıştır. Elde edilen sonuçların, doğruluk ve zaman başarımlarına göre, önerilen yöntemlerin seçimi için kriterler ve etkenler belirtilmiştir.

Öncelikle tez konusunu seçerken isteklerimi göz önünde bulundurup bana yardımcı olan ve tez konusu içeriği ile ilgili değerli bilgilerini paylaşan, çalışmamın her aşamasında önerileri ile beni yönlendiren danışman hocam Sayın Doç. Dr. Hasan Erdinç Koçer' e ve bugüne kadar her daim yanımda olan aileme ayrıca sevgili annem'e teşekkür eder, saygılarımı sunarım.

Mahmut Sami YASAK
KONYA - 2019

İÇİNDEKİLER

ÖZET	iv
ABSTRACT.....	v
ÖNSÖZ	vi
İÇİNDEKİLER	vii
SİMGELER VE KISALTMALAR	viii
1. GİRİŞ	1
1.1. XOR Şablon Eşleştirme Yöntemi	3
1.2. Karakter Tanınması Yöntemi.....	3
2. KAYNAK ARAŞTIRMASI	5
3. MATERYAL VE YÖNTEM.....	11
3.1. Görsellerin Toplanması.....	11
3.2. XOR Şablon Eşleştirme Yaklaşımı.....	14
3.3. MSER, SWT ve ESA ile Karakter Tanınması Yaklaşımı	15
4. ARAŞTIRMA BULGULARI VE TARTIŞMA	17
4.1. Dairesel Metal Cisimlerin Yerini Belirleme Problemi	17
4.1.1. Laplacian of Gaussian (LoG) Filtresi	17
4.1.2. Kenar Bulma Operatörü.....	18
4.1.3. Dairesel Hough Dönüşümü.....	20
4.2. Daugman's Rubber Sheet Modeli Çıkarımı.....	23
4.3. Ortalama Mutlak Sapma (OMS) ile Özellik Çıkarımı.....	25
4.4. XOR Şablon Eşleştirme	27
4.5. Maksimum Kararlı Ekstrimal Bölgeler (MSER) Uygulanması.....	28
4.6. Kontür Genişliği Dönüşümü (SWT) Gerçekleştirilmesi	29
4.7. Optik Karakter Tanınmasının (OCR) Gerçekleştirilmesi	31
4.8. Sharpen Filtresi Görüntülerinin ESA ile Doğrudan Sınıflandırılması.....	37
5. SONUÇLAR VE ÖNERİLER	41
5.1. SONUÇLAR.....	41
5.2. ÖNERİLER.....	42
KAYNAKLAR	43
ÖZGEÇMİŞ	47

SİMGELER VE KISALTMALAR

Simgeler

δ	: delta
f	: frekans
π	: pi
ψ	: psi
σ	: sigma
τ	: tav
θ	: theta

Kısaltmalar

CNN	: Convolutional Neural Networks – Evrişimli Sinir Ağları
DHD	: Dairesel Hough Dönüşümü - Circular Hough Transform
DRSM	: Daugman's Rubber Sheet Model
ESA	: Evrişimli Sinir Ağları
fps	: frame per second – saniye başına düşen görüntü sayısı
GF	: Gabor Filtresi
HSV	: Hue Saturation Value – Öz Doygunluk Parlaklık görüntü renk uzayı
LoG	: Laplacian of Gaussian
MSER	: Maximally Stable Extremal Regions – Maksimum Kararlı Ekstrimal Bölgeler
mAP	: mean Avarage Precision – Genel Hassasiyet Ortalaması
mm	: mili metre
ms	: mili saniye
OCR	: Optical Character Recognition – Optik Karakterlerin Tanınması
OMS	: Ortalama Mutlak Sapma
RGB	: Red Green Blue – Kırmızı Yeşil Mavi görüntü renk uzayı
s	: saniye
SF	: Sharpen Filtresi
SWT	: Stroke Width Transform – Kontür Genişliği Dönüşümü
SVM	: Support Vector Machine – Destek Vektör Makineleri
XOR	: XOR bit-düzeyi operatörü

1. GİRİŞ

Endüstriyel ürünlerin üretim, taşınma, ayırıştırma, sınıflandırma ve kalite kontrol safhalarında, yüksek doğruluk ve hız gibi başarımların parametrelerinin göz önünde bulunduğu sanayi toplumlarında, makineleşmeye ihtiyaç duyulmaktadır. Günümüz dünyasında makinelerin çözümleme kapasiteleri artıktıkça bir çok sektörün problemlerine cevap verebilmektedirler. İmalat sektöründe üretilen ürünlerin kalite kontrolleri için son yıllarda görüntü işleme tabanlı sistemler sıklıkla tercih edilmektedir. Makine görmesi tabanlı sistemler üreticilere hem hızlı hem de güvenilir çözüm sunmaktadır. Bu çalışmada da üzerinde oyma karakterler bulunan dairesel metal ürünlerin sınıflandırılmasına yönelik makine görmesi tabanlı bir çözüm sunulmuştur.

Tez çalışmasında kullanılan örnek ürünler, endüstriyel sanayi ortamında üretilen, üzerlerinde presle oyma işlemi gerçekleştirilmiş karakterler bulunan dairesel metal cisimlerden oluşmaktadır. Dairesel metal cisimlerin şekil itibari ile çapları farklı olup, yaklaşık olarak 27.85 mm ile 38 mm arasında değişmektedir. Yükseklikleri ise 6.07 mm ile 6.90 mm arasında değişirken, cisimlerin ayrıca etrafında bulunan flanş bölümü için yükseklik değerleri yaklaşık olarak 1.90 mm ile 3 mm arasında değişmektedir. Şekil 1.1’de çalışmada kullanılan dairesel metal cisimlerin örnek görselleri verilmiştir.



Şekil 1.1. Üzerlerine oyma işlemi gerçekleştirilmiş dairesel metal cisimlerin örnek görüntüleri

El ile yapılan ölçümler Euroricambi marka elektronik dijital kumpas ölçüm cihazı ile gerçekleştirilmiştir. Dairesel metal cisimler üzerinde, sınıflandırılabilirmeleri için marka, model ve seri numarası gibi bir çok amacı bulunan karakterler (harf, numara) ve semboller yer almaktadır. Bu karakter ve semboller, buldukları yüzey üzerine oyulmuş durumda bulunmaktadır. Oyma işlemi, dairese metal cisimler üzerine pres makinesi ile belli kalıpları/mühürleri bulunan karakterlerin basılması sonucu gerçekleştirilmiştir. Hassas olmayan bu işlem sonucunda, bazı karakterlerin oyma derinliği farklılık arz etmektedir. Dairesel metal cisimlerin örnekleme görüntüleri için, endüstriyel alan tarama kamera cihazı ve iki boyutlu lazer tarama cihazı kullanılmıştır. Endüstriyel alan tarama kamera cihazı ile elde edilen görseller için “kamera görsel sonuçları”, iki boyutlu lazer tarama cihazından alınan görüntüler için “profil çıkarım sonuçları” ifadeleri kullanılmıştır.

Bu çalışmada, oyma karakterlere sahip, endüstriyel dairese metal cisimlerin sınıflandırılma problemi için görüntü işleme algoritmaları kullanılarak iki farklı çözüm yolu sunulmaktadır. Önerilen ilk çözüm yöntemi XOR şablon eşleştirme yöntemidir. İkinci çözüm yolu ise, yapay zeka alanlarından derin öğrenme sınıflarından olan Evrişimli Sinir Ağları (ESA) algoritmasıdır.

Endüstriyel dairese metal ürünler üzerinde oyulmuş bir şekilde yer alan karakterlerin tanınmasını ve dairese metal cisimlerin sınıflandırılmasını konu alan bu çalışma presle şekil verilen ürünlerin kalite kontrolleri açısından büyük önem arz etmektedir. Bazı karakterlerin dairese bir şekilde yer alması sınıflandırmayı güçleştiren bir unsurdur. Günümüzde düz hat üzerinde yer alan karakterlerin tanınması otomobil plaka tanıma, levha, yazı, el yazısı, reklam tabelaları tanıma gibi alanlar ile yaygın olarak gerçekleştirilmektedir. Teze konu olan uygulamada ise, dairese bir hat üzerinde ve düz bir hat üzerinde bulunan karakterlerin tanınması birincil önem arz etmektedir. Ayrıca metal ürünler üzerinde oyma işlemi ile oluşturulan karakterler, karakterlerin boyutları hesaba katıldığında, iki boyutlu lazer tarayıcı cihazı ve alan tarama kamerası ile taranıp dijital ortama aktarılması ve sonuçlarının karşılaştırılması ikincil önem arz etmektedir. Elde edilecek bilgisayar uygulaması sonucunda ürünlerin kalite kontrol safhasında, sınıflandırılmasında veya ayrıştırılması gibi alanlarda kullanılabilir olması tezin önemini oluşturmaktadır.

Dairesel metal cisimlerin sınıflandırma problemi için önerilen her iki çözüm yöntemi için, örneklenecek ürünlerin görsellerine uygulanan tüm algoritmalar ve yöntemler, aşağıdaki iki madde içerisinde kısaca anlatılmaktadır.

1.1. XOR Şablon Eşleştirme Yöntemi

Görüntü işleme sürecinde nesne sınıflandırma problemi için uygulanan ilk yöntem XOR Şablon Eşleştirme (XOR Pattern Matching) yöntemidir. Bu yöntem için öncelikli olarak, kamera görsel sonuçlarına, Marr ve Hildreth (1980) tarafından bulunan Laplace of Gaussian (LoG) filtresinin uygulanması, Canny (1986) tarafından bulunan Canny kenar bulma algoritmasının uygulanması, görüntü piksel ilişkileri üzerine çalışmalarıyla ilk defa Hough (1962) tarafından ortaya atılan cisim kenar piksel ilişkileri, ve teoremin dairesel cisimler için Ylajaaski ve Kiryati (1994); Lam ve Yuen (1996); Smereka ve Duleba (2008) tarafından geliştirilen yöntemlerince Dairesel Hough Dönüşümü (DHD) algoritmasının uygulanması ile dairesel metal cisimlerin görüntü içerisindeki yerinin tespit edilmesi problemi çözülmüştür. Elde edilen görüntünün Daugman (1997) tarafından geliştirilen bir teknik ile Daugman's Rubber Sheet Model (DRSM)' inin çıkarılması, Gabor (1946) tarafından bulunan Gabor Filtresinin (GF) uygulanması ve sonucunda faz görüntüsünün kullanılması, görüntünün Ortalama Mutlak Sapma (OMS) çıkarımı şeklinde, görüntülerin normalize edilmesi ve özellik çıkarımları gerçekleştirilmiştir. Özellik çıkarımı elde edilen görüntüler belirli matris işlemlerine tabi tutulur. Bu matris işlemleri sırasıyla, görüntünün eşiklenerek ikili görüntüye dönüşümü, matris transpozunun alınması ve görüntüyü 99x750 boyutundan 1x74250 şeklinde yatay hizada uzatılması işlemleri gerçekleştirilmiştir. Elde edilen bu matix uzayı her bir örneklenecek görsel için daha sonra eşleştirilmek üzere veri tabanına aktarılarak, öncelikli işlemler tamamlanır. Nesne sınıflandırma problemine alınan her bir yeni ham görüntü için, öncelikli işlemler tekrarlanarak, veri tabanındaki her bir matris uzayı ile XOR operatörü işlemine sokulur. XOR işlemi sonrasında elde edilen mantıksal "0" ve "1" değerlerinin oranı bulunur. "0" ikili (binary) değerinin görüntü içerisindeki oranı şablon eşleştirmenin %' lik oranını belirlemektedir. Şablon eşleştirme oranı en yüksek olan, veri tabanındaki şablon ismi ile eşleşerek, nesne sınıflandırma problemi çözülmüştür.

1.2. Karakter Tanınması Yöntemi

Nesne sınıflandırma problemlerin çözümü için gerçekleştirilen yaklaşımlardan biri olan ve ilk tanımı D'Albe (1920) tarafından konulan ve icat edilen Optophone, daha sonra günümüz karakter tanınması için kullanılan Optik Karakter Tanınması (Optical Character Recognition: OCR) halini almıştır. Nesne sınıflandırma problemine bir başka

yaklaşım olan karakter tanınması yöntemi ile dairesel metal cisimlerin sınıflandırma problemi, aşağıdaki paragrafta belirtilen, sırasıyla uygulanan algoritmalarca çözülmüştür.

Görüntü işleme sürecinde uygulanan ikinci yöntem olan OCR yöntemi için ön işlem yöntemleri sırasıyla LoG filtresinin uygulanması, Canny kenar bulma algoritmasının uygulanması ve son olarak DHD algoritması uygulanmasıdır. Sıralı işlemler uygulanarak dairesel metal cisimlerin görüntü içerisindeki buldukları yeri belirleme problemi çözülmüştür. Dairesel metal yüzeyler üzerindeki karakterler, cisimlerin tam ortasında yer almakta ve cisimlerin belirli bir yüksekliğinde bulunmaktadır. Bu bölümde yer alan karakterler için karakter bölgesi ifadesi kullanılacaktır. Ayrıca, karakter bölgesinin bulunması için kamera görsel sonuçları bu kriterlere göre alınmıştır. Yeri belirlenen cisimlerin, karakter bölgesinin bulunması için DHD algoritması uygulamasından elde edilen sonuçlara göre 3 farklı daire çapı elde edilmiştir. En küçük veya en içte kalan dairenin içinde kalan bölge, dairesel metal yüzeylerin ön karakter bölgesi olarak tanımlanmakta ve ön karakter bölgesi görüntü içerisinden kesilip alınmıştır. Ön karakter bölgesi çıkarılan görsellerin karakter alanlarının bulunması (karakter segmentasyonu) için sırasıyla, kontrast genişletme işlemlerinden olan Sharpen Filtresinin (SF) uygulanması, Mikolajczyk ve ark. (2005) tarafından geliştirilen Maximally Stable Extremal Regions (MSER) algoritmasının uygulanması ve Epshtein ve ark. (2010) tarafından geliştirilen Stroke Width Transform (SWT) algoritması uygulanmıştır. MSER algoritması görüntü içerisindeki karakter kontürlerini bulurken, SWT algoritması karakter kontür genişliğini hesaplayıp, MSER algoritmasının bulunduğu bölgeler içerisindeki muhtemel karakter olmayan bölgeleri elemektedir. Ayrıca, karakterlerin dairesel, ters veya polar düzlemi hizalarında bulunması durumları için uygulanan MSER ve SWT algoritmaları başarılı sonuçlar vermiştir. Karakter segmentasyonu sonucu elde edilen karakter görselleri OCR işlemi için hazır hale getirilmiştir. OCR işlemi için, Fukushima (1980) tarafından ortaya atılan hiyerarşik çok katmanlı derin sinir ağları (Neocognitron) sınıflarından biri olarak LeCun ve ark. (1998); Hinton ve ark. (2006) tarafından geliştirilen ESA modeli tasarlanarak, daha sonra elde edilen görüntülerin çoğaltılması ve Cohen ve ark. (2017) tarafından oluşturulan EMNIST veri setlerince model eğitilmiş ve karakter görsellerine uygulanmıştır. Modelin çıkış katmanından (softmax layer - dense layer - classification layer) elde edilen sonuçlar, karakterleri tanımlamış ve etiketlemiştir. Dairesel metal cisimlerin nesne sınıflandırma problemi için karakter tanınması çözümü, bu çalışmanın ikinci yöntemini oluşturmaktadır.

2. KAYNAK ARAŞTIRMASI

Bu bölümde nesne sınıflandırma problemi, karakter tanınması, karakter segmentasyonu problemi ve XOR şablon eşleştirme yöntemi ile bu problemlere ve uygulanan çözüm yöntemlerine ilişkin literatür taraması sonuçları sunulmuştur.

Yapay zekanın derin öğrenme sınıflarından ESA algoritması yaygın olarak kullanılmaktadır. Nesne sınıflandırma problemi için sıklıkla kullanılan Redmon ve Farhadi (2017) tarafından geliştirilen YOLOv2 (YOLO-You Only Look Once) yada diğer ismi ile YOLO9000 algoritması 9000 sınıf nesne sınıflandırabilirken, nesnelerin gerçek zamanlı olarak değişen algılama çerçevelerini VOC metrik sınıflandırmasında %73.4 mAP (mean Average Precision - mean Ortalama Hassasiyet) oranında tahmin edebilmektedir. Bu işlemi, 512x512 çözünürlüğe sahip, 40 fps tazemeleme hızlı gerçek zamanlı video görüntüler üzerindeki nesnelerin, tahmin edilmesi olarak gerçekleştirmiştir.

Tekrar Redmon ve Farhadi (2019) tarafından geliştirilen ve daha gelişmiş olan YOLOv3 algoritması, COCO veri setlerini kullanarak, COCO mAP metrik sınıflandırmasında %51.5 mAP oranı ile 22 ms sürede 320x320 çözünürlüğe sahip görüntüler üzerinde nesne sınıflandırma problemini çözmüştür.

Sınıflandırma için hız ve doğruluk oranları yüksek olan YOLO algoritmaları farklı alanlardaki problemlerin çözümü için, kendi veri setine sahip görüntüler ile eğitilmektedir. Örneğin; Tian ve ark. (2019) tarafından yapılan bir çalışmada, meyve bahçelerinde elmaların gerçek zamanlı olgunluk sınıflandırılması için kullanılan YOLOv3 algoritmasına Dense katmanları eklenerek, %81.7 doğruluk oranına sahip 3000x3000 çözünürlüğünde görüntüler içerisindeki elma olgunluğu tahmin edilmiştir.

Huizhong ve ark. (2011) tarafından gerçekleştirilen örnek bir çalışmada, karakter bölgesinin tespiti problemi için, kenarları-geliştirilmiş MSER algoritması, geometrik ve kontur bilgileri kullanılarak, doğal resimler üzerindeki karakterlerin bölgesel tespitini gerçekleştirmiştir. Bu çalışmada, Canny kenar bulma operatörü ve MSER algoritmasının kombinasyonu kullanılarak, ICDAR 2003 ve 2005 veri setleri üzerinde oldukça küçük karakterli ve bulanık karakterli görsellerin karakter bölgesi ve segmentasyonu problemini, %73 doğruluk oranında, %60 en iyi eşleşme oranında, hassaslık ve en iyi eşleşme oranının kombinasyonu olan %66 f metrikleri ile çözmüştür.

Gonzalez ve ark. (2012) tarafından gerçekleştirilen başka bir çalışmada, karakter tanıma problemi için, farklı font, yazı tipi, yazı kalınlığı, yazı geometrisi, yazı renkleri,

yazı dokusu, yazı boyutu, yazı bulanıklığı gibi zorlu koşullar içeren görseller için çözüm getirmiştir. MSER ve adaptif eşikleme yöntemlerinin birleşimi sonucu karakter segmentasyonu için çözüm geliştirilerek, görsel içerisinde karakter bölgesinden farklı olarak metin dışı şekillerin veya nesnelerin tespit edilip, görsel dışına çıkarılması gerçekleştirilmiştir.

Metin dışı nesnelerin ve metin karakterlerinin ayırt edilebilmesi için Chapelle ve ark. (1999) tarafından geliştirilen Destek Vektör Makinesi'ne dayalı sınıflandırıcılardan olan, Hanif ve Prevost (2009) tarafından geliştirilen Mean Difference Feature (MDF) yöntemi, standart sapma algoritması ve Dalal ve Triggs (2005) tarafından geliştirilen Histograms of Oriented Gradient (HOG) yöntemi kullanılmıştır. ICDAR 2003 veri setleri üzerinde çalışmanın uygulanarak, %81 doğruluk oranında, %57 en iyi eşleşme oranında, hassaslık ve en iyi eşleşme oranının kombinasyonu olan %67 f metrik oranları ile karakter bölgesi tespiti ve karakter segmentasyonu problemini çözmüştür.

Li ve Lu (2012) tarafından gerçekleştirilen karakter segmentasyonu için kontrastı-geliştirilmiş MSER algoritması uygulanıp, sonrasında SWT algoritmasının uygulaması ile karakter bölgesi tespiti problemi için yapılan başka bir çalışmada, geometrik kısıtlamalar ile metin dışı olan nesnelerin tespiti ve görsel içerisinde elenmesi yöntemleri sunulmuştur. İlk olarak, Matas ve ark. (2004) tarafından geliştirilen temel MSER algoritması ile elde edilen karakter bölgesinin bulanık olma durumları için, geliştirilen kontrastı-geliştirilmiş MSER algoritması ile bulanık karakterlere sahip görüntülerin karakter bölgesinin tespiti gerçekleştirilmiştir. İkinci olarak, Zhao ve ark. (2011) çalışmasında kullanılan en boy oranı, 0.3 mm - 3 mm (3 - 30 pixel) arası karakter boyutları ve uygunluk değerlerine göre belirlenen parametreler kullanılarak geometrik kısıtlama işlemi sonrasında metin dışı nesnelere görselden çıkarılmıştır. Son olarak, metinlerin hizasını analiz etmek için metinlerin kontur genişliğini kullanan Shivakumara ve ark. (2011)'nin algoritması uygulanıp, karakterlerin iskelet'i çıkarılarak, karakter bölgesi tespit problemi çözülmüştür. ICDAR 2003 veri setleri üzerinde çalışmanın uygulanarak, %59 doğruluk oranında, %59 en iyi eşleşme oranında, hassaslık ve en iyi eşleşme oranının kombinasyonu olan %59 f metrik oranları ile karakter bölgesi tespiti ve karakter segmentasyonu problemleri çözülmüştür.

Neumann ve Matas (2012) tarafından gerçekleştirilen gerçek zamanlı ve düşük kontrastlı görüntüler için karakter bölgesinin tespiti ve karakter tanınması problemleri için yapılan başka bir çalışmada, düşük kontrastlı alanların, karakter bölgesinin tespiti için taranmasıdır. Bu yöntemle zorlu bölgeler (Extremal Region - ER) adı verilerek, bölge

içerisindeki karakterlerin dış piksellerinin RGB ve HSI değerleri, bölgenin kendisinin RGB ve HSI değerlerinden daha yüksek olduğu yaklaşımı getirilmiştir. Bu aşamada, iki çeşit sınıflandırma kriteri belirtilmiştir. İlk olarak, verimli bulunan eşikdeğeri ile elde edilen binary görüntünün, 2x2 piksel desenli Pratt (2001) tarafından ortaya konan Euler sayıcısı ile sayılarak, Schapire ve Singer (1999) tarafından geliştirilen, en boy oranı ve kompaktlık gibi parametreler kullanan karar verme ağacı/sınıflandırıcısından biri olan Real AdaBoost sınıflandırıcısı kullanılarak görüntünün topolojik özelliği elde edilir. İkinci olarak, elde edilen topolojik özellikli görsel SVM' ye dayılı sınıflandırıcılardan olan, Müller ve ark. (2001) tarafından geliştirilen Radial Basis Function (RBF) yöntemi uygulanarak ikincil topolojik özellikli görsel elde edilmiştir. İkincil topolojik özelliği tespit edilen görselin RGB ve HSI sonuçları ile karşılaştırılması sonucu, düşük kontrastlı bölgeler de karakter bölgesi çıkarımı problemi çözülmüştür. ICDAR 2003 veri setleri üzerinde çalışmanın uygulanarak, %73.1 doğruluk oranında, %64.7 en iyi eşleşme oranında, hassaslık ve en iyi eşleşme oranının kombinasyonu olan %68.7 f metrik oranları ile karakter bölgesi tespiti ve karakter segmentasyonu problemleri çözülmüştür.

Zhou ve ark. (2017) tarafından gerçekleştirilen başka bir çalışmada, nesne sınıflandırma problemi için son derece kullanışlı olan ESA derin öğrenme algoritması başarılı bir şekilde optik karakterler ve el yazısı gibi karakterlerin tanımlama problemi içinde kullanılmaktadır. Örnek bir çalışmada, karakterlerin görsel içerisindeki bölgesini bulmak için geliştirilen derin öğrenme sınıflarından biri olan ve Long ve ark. (2015) tarafından geliştirilen Fully Convolution Network (FCN) modeli PVANet2x isimli model tanıtılmaktadır. Modelin tasarımı için, görüntü içerisindeki karakter bölgelerinin büyüklükleri ve karakter bölgelerinin geometri parametreler değerlendirilmiş, 4 Conv katmanlı + 3 Unpooling katmanlı + FC katmanlı + 4 sınıflandırma değeri veren model oluşturulmuştur. Modelin eğitimi için ICDAR 2015 veri seti kullanılarak, yine ICDAR metrik sınıflandırmasında % 83.27 doğruluk oranına, %78.33 en iyi eşleşme oranında, hassaslık ve en iyi eşleşme oranının kombinasyonu olan %80.72 f metrik oranları ile karakter bölgesi tespiti problemini çözmüştür.

Xie. ve ark. (2018) çalışmasında, karakter bölgesi tespiti problemi için, ESA sınıflarından biri olan ve He ve ark. (2017) tarafından geliştirilen güçlendirilmiş-ESA algoritmasından esinlenen ve Ma ve ark. (2018) tarafından oluşturulan Feateare Pyramid Network (FPN) algoritmasını dahada geliştirerek Supervised Pyramid Context Network (SPCNET) algoritması tanıtılmaktadır. Modelin yapısı çok katmanlı olup sırasıyla, FPN yapısı katmanlı + karakter bölgesi ön sınıflandırma için çalışmalarında geliştirdikleri

metin bağlamı sınıflandırıcı modülü + mask RCNN modeli şeklinden oluşmaktadır. Mask RCNN modeli için ResNet50 modeli kullanılmış olup, ICDAR 2015 veri seti kullanılarak, ICDAR metrik sınıflandırmasında % 88.7 doğruluk oranına, %85.8 en iyi eşleşme oranında, hassaslık ve en iyi eşleşme oranınının kombinasyonu olan %87.2 f metrik oranları ile karakter bölgesi tespiti problemini çözmüştür .

Karakter bölgesi tespiti ve karakter segmentasyonu problemleri için MSER ve SWT algoritmalarının sıklıkla kullanıldığı yukarıdaki çalışmalarda anlatılmıştır. Bu tezde önerilen ikinci yöntem olan OCR yönteminin ön aşamasında MSER ve SWT algoritmaları kullanılacaktır.

Karakter tanınması problemi için, el yazısı veri setlerini kullanan Federico ve ark. (2009) tarafından gerçekleştirilen bir çalışmada, Vintsyuk (1968) tarafından tanıtılan Dinamik Zaman Atlamalı (Dynamic Time Warping - DTW) karşılaştırılmasına dayanan algoritma ile karakterlerin şablonu çıkarılarak, şablon eşleştirme tekniği ile %0.60 hata oranına sahip karakter tanınması gerçekleştirmişlerdir. Görselin şablon eşleştirme için veri tabanına yüklenmeden önce, DTW algoritması ile görüntü pikselleri minimum maliyet yolu optimizasyonu kullanılarak gezilip, piksel ilişkisi ağırlık oranları çıkartılır. Piksel ilişkisi ağırlık oranları, piksel ilişkisi (en yakın komşu pikseller) belirlenen pikseller arası öklid mesafesinin karesini alarak hesaplanır. Bu ağırlık oranları ile görselin şablonu çıkarılarak veri tabanına kaydedilir. Karakter tanınması gerçekleştirilecek görüntüye aynı işlemler uygulanarak, en yakın komşu pikseller kuralı ile eşleştirme uygulanır ve etiketlenmiş karakter tespit edilerek karakter tanınması problemi çözülmüştür.

Radzi ve Khalil-Hani (2011) tarafından gerçekleştirilen başka bir çalışmada, derin öğrenmenin ESA sınıfı kullanılarak araç plaka tanıma sistemi için karakter segmentasyonu gerçekleştirilmiş karakterlerin, tanınması işlemi çalışılmıştır. ESA modelinin eğitilmesi için, 22x12 çözünürlüğünde 750 adet eğitim için, 434 adet ise doğrulama görseli kullanmışlardır. Tasarladıkları ESA modeli sırasıyla, 5x5 filtre matrisine sahip 6 adet filtre sayılı Convolution katmanı, 2x2 filtre matrisine sahip 6 adet filtre sayılı Subsampling katmanı, 5x5 filtre matrisine sahip 16 adet filtre sayılı Convolution katmanı, 120 nöronluk Fully Connected katmanı ve 33 nöronluk Fully Connected katmanından oluşmaktadır. Bu şekilde ESA çıkışını, 33 sınıflandırma indeksi verecek şekilde ayarlamışlardır. ESA modeli eğitimi için aktivasyon fonksiyonu ‘tanh’ olarak seçmişlerdir. Eğitim sonunda, % 1.21 hata oranı veya %98.79 doğruluk oranı yakalanarak, karakter tanınması problemi çözülmüştür.

Bu tezde önerilen ikinci yöntem olan karakter tanınması için ESA kullanılmasına nesne sınıflandırma problemi içerisinde anlatılan literatür araştırmasında belirtilmiştir. Yalnız, ESA modeli için yeniden tasarlanması belirtilmiş ve Radzi ve Khalil-Hani (2011) çalışmasında tasarlanan model parametrelerinden faydalanılacaktır.

Nesne sınıflandırma problemi için görüntü işleme yöntemlerinin gerçekleştirdiği bir başka yaklaşım olan şablon eşleştirme yöntemi, özellik çıkarımı gerçekleştirilmiş ana görüntülerin şablon olarak veri tabanına kaydedilmesi ve eşleştirilmek için gelen resimler ile işleme sokularak, eşleşme yüzdelerine göre karar veren temellerden oluşmaktadır.

Bala ve Tajinder (2016) tarafından gerçekleştirilen örnek bir çalışmada, nesne sınıflandırma problemini veya içerik tabanlı görüntü sınıflandırma (Content-Based Image Retrieval - CBIR) problemini (Subrahmanyam ve ark., 2013), özellik çıkarımı gerçekleştirilen Smith (1978) tarafından tanımlanan HSV histogramı ve lokal texton XOR pattern (LTxXORP) adı ile adlandırdığı bir öneri ile çözmüştür. HSV bilgileri elde edilen görüntülere, belirlenen yapısal bileşenli matris' i Heikkila ve ark. (2009) tarafından geliştirilen Scale Invariant Feature Transform (SIFT) yöntemi ile gezdirilerek Ojala ve ark. (1996) tarafından ortaya konulan Local Binary Pattern (LBP) yönteminin geliştirilmesi sonucu LTxXORP değerinin elde edilmesi sağlanmıştır. Elde edilen LTxXORP değerleri tüm görüntüye uygulanarak, görüntü ikili "0" ve "1" değerlerine indirgenip karşılaştırılmak üzere etiketlenerek veri tabanına kaydedilir. İçerik tabanlı görüntü sınıflandırma için kullanılacak görsellere aynı işlem uygulanarak, tüm görüntülerin bulunduğu veri tabanındaki ikili değerlere XOR operatörü işlemi uygulanır ve eşleşme yüzdesi üretilir. Eşleşme yüzdesi ile benzer görseller sınıflandırılmış olunur. Bu şekilde şablon eşleştirme yöntemi ile nesne sınıflandırma veya CBIR problemi çözülmüştür.

Nesne sınıflandırma problemi içerisinde yer alan insan ayırımı veya insan kimliğini belirleme gibi problemler için çözüm getiren Chawla ve Oberoi (2011) çalışmasında, insan gözü iris katmanı görseller kullanılarak, sırasıyla, görüntünün DRSM modeli çıkarılarak, modelin GF ile özellik çıkarımı gerçekleştirilmiştir. Elde edilen görüntü binary görüntüye dönüştürülerek, veritabanındaki görüntüler ile XOR operatörü işlemine sokulmuştur. Daugman (2003) tarafından geliştirilen ve 9.060.003 iris karşılaştırması sonucu elde edilen, ortalama, standart sapma, maksimum ve minimum Hamming Mesafe (Hamming Distance: HD) verileri ile korelasyon veya eşleşme ilişkisi kurulmuştur. CASIA verileri kullanılan bu çalışmada, ortalama eşleşme süresi 133.7

saniye ve hamming distance oranı 0.3486 oranları ile insan kimliği belirleme problemi çözülmüştür.

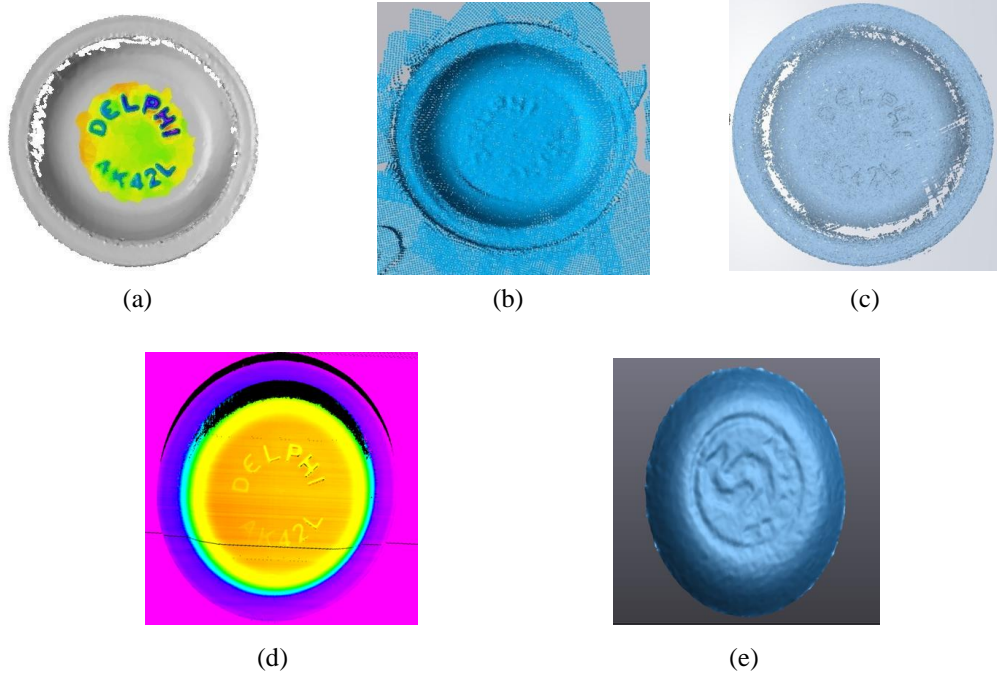
Daugman' nın kullandığı iris görüntülerin dairesel oluşu, bu tezde kullanılan örnekleme görüntülerin de dairesel oluşu düşünüldüğünde, önerilen XOR şablon eşleştirme yöntemi için özellik çıkarımı öncesinde kullanılmasına ilham kaynağı olmuştur. Ayrıca Daugman (2003); Bala ve Tajinder (2016) çalışmalarında şablon eşleştirme için kullanılan XOR operatörü, bu tez çalışmasında kullanılacaktır.



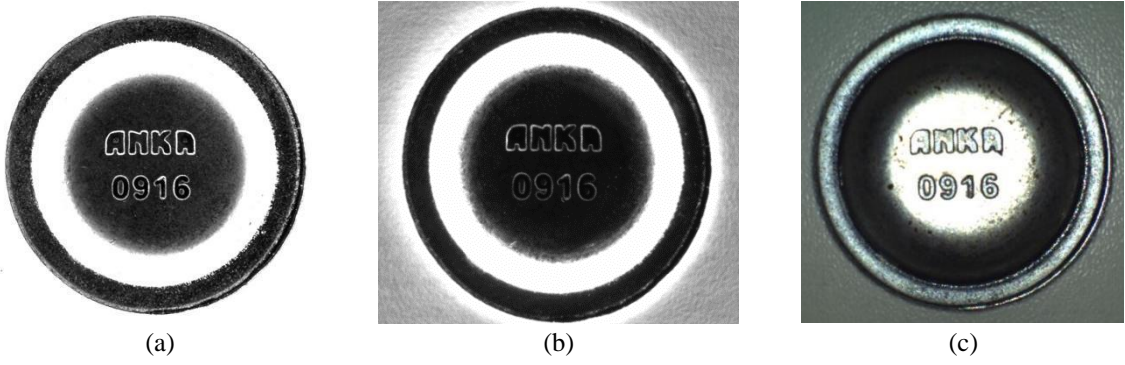
3. MATERYAL VE YÖNTEM

3.1. Görsellerin Toplanması

Teze konu olan dairesel metal cisimlerin görselleştirilmesi için iki farklı cihaz (Lazer tarayıcı ve Alan tarama kamerası) kullanılmıştır. İki farklı cihaz kullanılmasının sebebi, dairesel metal cisimler üzerindeki karakterlerin, yazı-karakteri (font) ve karakterlerin boyutlarının farklı oluşudur. Aynı zamanda iki farklı cihazdan elde edilen sonuçlar değerlendirilmiş ve oyma karakterlere sahip dairesel metal cisimlerin sınıflandırma problemi için verimli ve ihtiyaçları karşılayan cihazın tespiti yapılmıştır. Dairesel metal cisimlerin görselleştirilmesi için iki boyutlu lazer tarama cihazı ve endüstriyel alan tarama kamera cihazı kullanılmıştır. Profil çıkarım sonuçlarının ve kamera görsel sonuçlarının alınması için farklı marka ve modelde cihazlar denenmiştir. İki boyutlu lazer tarama cihazından alınan en iyi profil çıkarımı weCube programı ile Wenglor marka MLWL2 model cihazından alınmıştır. Kamera görseli içinse In Sight Explorer programı ile Cortex marka In Sight 7010 model cihazından en iyi görüntü alınmıştır. Farklı marka ve modellerden alınan profil çıkarım sonuçları Şekil 3.1.' de gösterilmektedir. Benzer şekilde farklı marka ve modellerden alınan kamera görsel sonuçları da Şekil 3.2.' de gösterilmektedir.



Şekil 3.1. (a) Faro marka lazer tarayıcı profil çıkarım sonucu, (b) Leica marka lazer tarayıcı profil çıkarım sonucu, (c) Nikon marka lazer tarayıcı profil çıkarım sonucu, (d) Wenglor marka lazer tarayıcı profil çıkarım sonucu, (e) Creaform marka lazer tarayıcı profil çıkarım sonucu



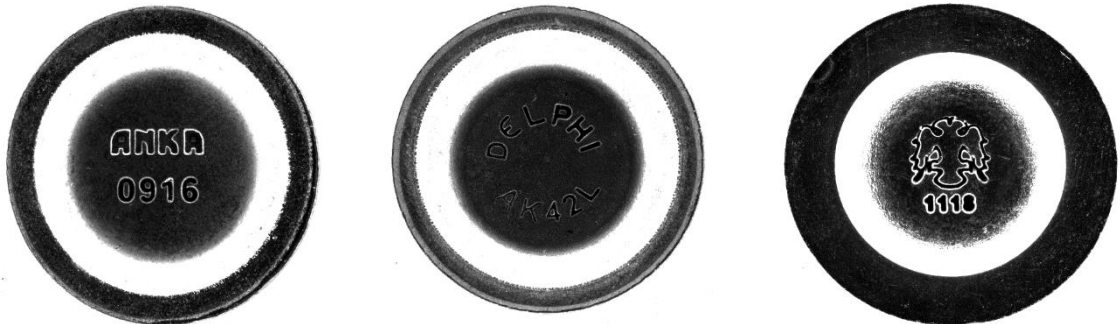
Şekil 3.2. (a) Cognex marka kamera görsel sonucu. (b) Sick marka kamera görsel sonucu, (c) Wenglor marka kamera görsel sonucu

Şekil 3.1.' de görüldüğü üzere en iyi profil çıkarım sonucu Wenglor markasından alınan (Şekil 3.1.d) görsel seçilmiştir. Şekil 3.2.' den görüldüğü üzere en iyi kamera görsel sonucu Cognex markasından alınan (Şekil 3.2.a) görsel seçilmiştir.

Kamera görsel sonuçları için bir başka parametre ise en iyi kamera markası seçildikten sonra, kamera için ışık kaynağının belirlenmesi olmuştur. Dom ışık kaynağı ve Dark-Field ışık kaynağı olarak iki farklı ışık kaynağı kullanılmıştır. Dom ışık kaynağı ile alınan kamera görsel sonuçları Şekil 3.3.' te, Dark-Field ışık kaynağı ile alınan kamera görsel sonuçları ise Şekil 3.4.' de verilmiştir.

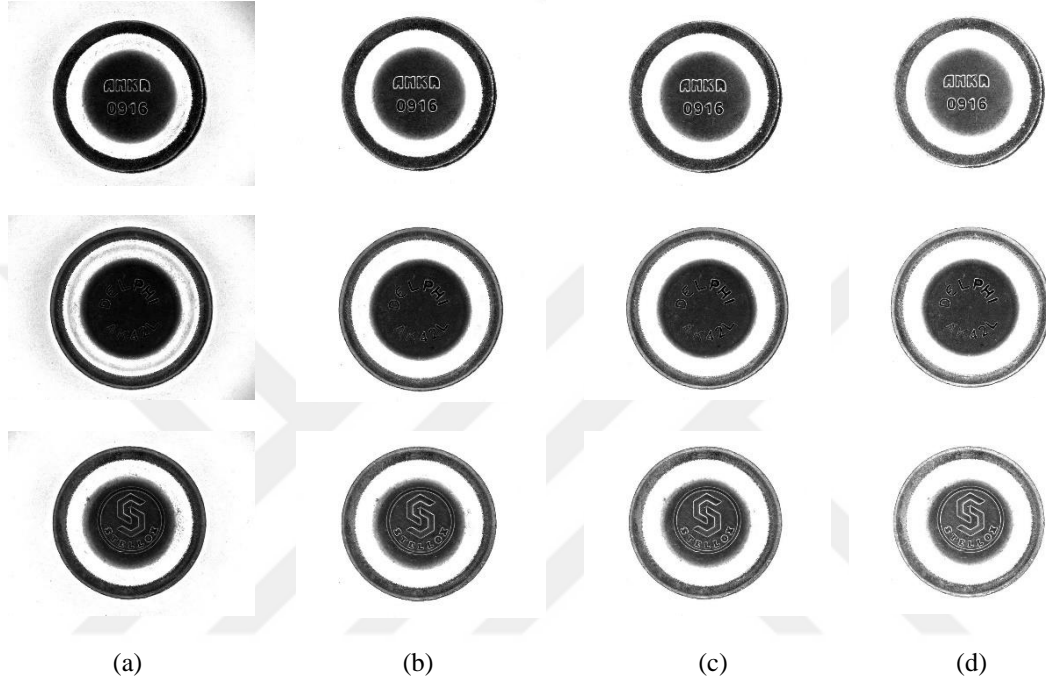


Şekil 3.3. Dom ışık kaynağı kullanılarak pozlama süresi 8ms olan kamera görsel sonuçları



Şekil 3.4. Dark-field ışık kaynağı kullanılarak pozlama süresi 10 ms olan kamera görsel sonuçları

Işık kaynağı uygulama sonuçlarına göre, karakter bölgesi belirleme probleminin çözümünde dairesel metal cisimlerin daire kenar konturları daha net bir şekilde görülen Dark-Field ışık kaynağı seçilmiştir. Kamera görsel sonuçları için bir başka seçim parametresi ise, ışık kaynağının pozlama süresidir. Şekil 3.5.' te farklı pozlama süreleri ile alınan örnek kamera görüntüleri verilmiştir.



Şekil 3.5. Dark-field ışık kaynağının (a) pozlama süresi 7 ms, (b) pozlama süresi 10 ms, (c) pozlama süresi 12 ms, (d) pozlama süresi 15 ms olarak ayarlanıp elde edilen kamera görsel sonuçları

Dairesel metal cisimler içerisindeki karakter bölgesinin korunması ve daire kontür bilgilerinin korunması faktörleri ile test edilen sonuçlara göre pozlama süresi 10 ms olan görüntüler, dairesel metal cisimlerin sınıflandırılması için kullanılmıştır. Işık şiddetinin 7 ms ve daha az olduğu durumlarda dairesel metal cisimlerin kenar kontürleri daha iyi görünmesine karşı, karakterlerin görüntüsü oldukça bulanık çıkmakta, karakter bölgesi tespiti problemi başarısız olmaktadır. Işık şiddetinin 15 ms ve üstü durumlarda ise karakter bölgesinin daha iyi görünmesine karşı, dairesel metal cisimlerin kenar kontürleri daha bulanık çıkmaktadır. Kenar kontürlerinin tespit edilmesindeki asıl amaç MSER algoritmasının uygulanmasında verimli bir sonuç alınması için karakter bölgesinin oldukça belirgin olması gerekmektedir. Dolayısıyla, kamera görsel sonuçlarında elde edilen görsellerin Şekil 3.5' te görüldüğü üzere en iç dairesinde bulunan karakter bölgesi kesilip alınmalı, karakter dışı olan bölgeler görsel dışına çıkarılarak elenmelidir. Karar verilen 10 ms pozlama süresi sonrasında, kamera ayarları olarak, diyafram değeri f/16

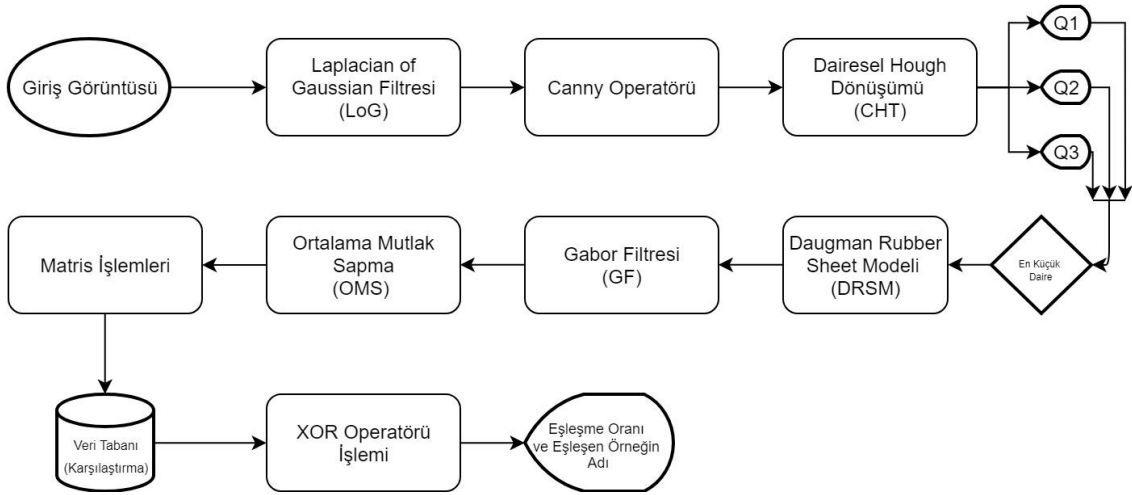
olarak ayarlanan, 35 mm' lik Cmont marka lens ile 800x600 çözünürlüğünde giri seviye görüntüler alınmıştır.

Şekil 3.1.' de profil çıkarım sonuçları görüldüğü üzere, dairesel metal cisimlerin ve üzerinde bulunan karakterlerin boyutlarından dolayı iki boyutlu lazer tarayıcı cihazının kullanılması, tam daire elde edilememesi, karakterlerin cisim yüzey geçişlerindeki yumuşaklık ve karakterlerin oldukça gürültülü görselleştirilmesi sebepleri ile görüntü işleme tekniklerince yetersiz kalacağından, ayrıca lazer tarayıcı cihazının tarama gerçekleştirirken gerçekleştirdiği gürültü sonucu profil çıkarım sonuçlarının etkilenmesi sebebi ile sınıflandırma problemi için örnekleme görseli olarak kullanılmamasına karar verilmiştir.

Çalışma kapsamında gerçekleştirilen tüm işlemler, Intel i7 4720HQ 2.6GHz CPU, 16GB 2133MHz RAM donanımlarına sahip bir bilgisayarda yapılmıştır. ESA modelin eğitilmesi için ise NVidia GTX960M 2GB RAM GPU donanımı kullanılmıştır. Çizelgelerde belirtilen, uygulanan algoritmaların işlem süreleri yukarıdaki donanımlar ile elde edilmiş olup, çalışmanın programlama aşaması Matlab R2018a programında geliştirilmiştir.

3.2. XOR Şablon Eşleştirme Yaklaşımı

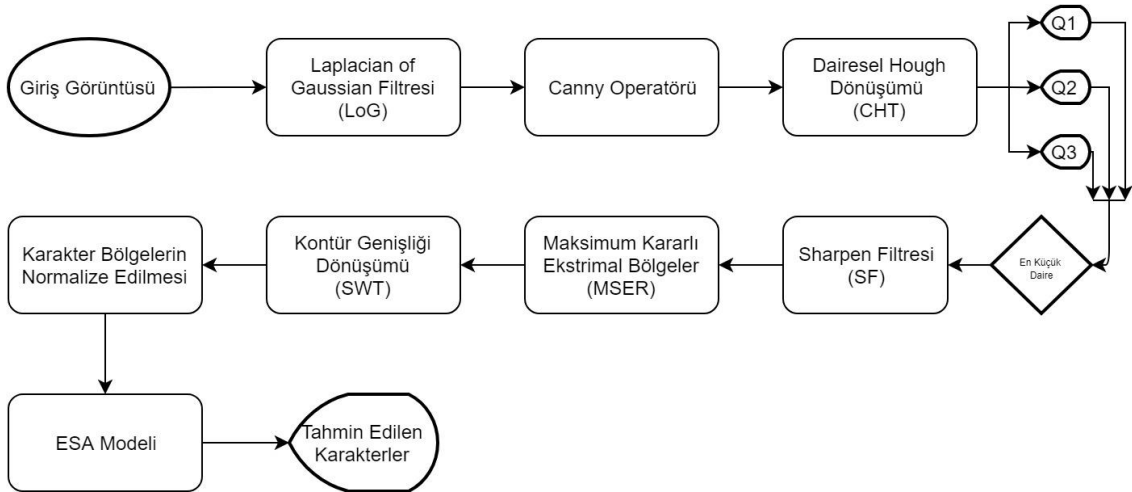
XOR şablon eşleştirme yaklaşımı, dairesel metal cisimlerin sınıflandırılması için uygulanan ilk yöntemdir. Görüntü işleme prosedürleri için seçilen 10 ms pozlama süreli kamera görsel sonuçları kullanılmıştır. XOR şablon eşleştirme yaklaşımı kısaca, giriş görüntüsü olarak adlandırılan örneklenecek görüntü ile özellik çıkarımı gerçekleştirilmiş ve veritabanına yüklenmiş görüntü xor operatörü işlemine sokulur. İşlem sonucu elde edilen '0' ve '1' ikili değerlerinden, '0' değerinin, işlem sonucu görüntünün boyutuna oranı, eşleşme oranı olarak tanımlanır. Eşleşme oranı en yüksek olan etiketlenmiş örnek adı veritabanından çağrılarak eşleşme gerçekleştirilir. XOR operatörü işlemi gerçekleştirilmeden önce, bir dizi görüntü işleme algoritmaları uygulanmıştır. Şekil 3.6.' da uygulanan algoritmaların akış diagramı verilmiştir. Akış diyagramından da görüldüğü üzere ham görüntüye öncelikle LOG filtresi uygulanarak kenar bulma işleminin doğruluğu artırılmıştır. Canny kenar bulma işlemi sonrasında dairesel Hough dönüşümü uygulanarak en dış, orta ve iç daire sınır çizgileri çizilmiştir. İç daire sınırları içinde kalan kısımdaki görüntüye DRSM uygulanarak polar-kartezyen dönüşüm yapılmıştır. Ardından Gabor filtresi ve OMS uygulanmış ve özellik vektörü çıkarılmıştır. Son aşamada ise XOR operatörü ile eşleştirme yapılmıştır.



Şekil 3.6. XOR şablon eşleştirme yöntemi için uygulanan tüm algoritmaların akış diagramı

3.3. MSER, SWT ve ESA ile Karakter Tanınması Yaklaşımı

Dairesel metal cisimlerin sınıflandırılması problemi için önerilen ikinci yöntem, cisimler üzerindeki karakterlerin tanınarak sınıflandırılmasının gerçekleştirilmesidir. Bu yöntem içinde, görüntü işleme prosedürleri için seçilen 10 ms pozlama süreli kamera görsel sonuçları kullanılmıştır. Şekil 3.7.'de karakter tanınması probleminin çözümü uygulanan algoritmaların akış diagramı verilmiş olup, her bir algoritma için gerekçe, süreç ve sonuçları aşağıda oluşturulan alt başlıklarca anlatılmaktadır.



Şekil 3.7. Karakter tanınması yöntemi için uygulanan tüm algoritmaların akış diagramı

Karakter tanınması problemi için karakter bölgesinin üzerinde bulunduğu dairesel metal cisimlerin görüntü içerisindeki yerinin tespit edilmesi ve karakter tahmini doğruluk oranının yüksek olabilmesi için karakter bölgesi olabildiğince bulunduğu görüntüden ayrılması gerekmektedir. Dolayısıyla, XOR şablon eşleştirme bölümünde gerçekleştirilen, dairesel metal cisimlerin yerini belirleme problemi için kullanılan LoG

filtresi, Canny operatörü ve DHD algoritması, daire çaplarının bulunup en iç daire içerisinde bulunan karakter bölgesinin çıkarımı işlemi tekrarlanmaktadır. Daha sonra elde edilen karakter ön bölgesi bir sonraki aşama olan MSER algoritması uygulaması için hazır hale getirilmiştir. Karakter tanıma işleminde MSER ve SWT algoritmaları ile karakter segmentasyonu yapılmıştır. Segmente edilen karakter resimleri, ESA modeli ile sınıflandırılarak karakterler tanınmıştır.



4. ARAŞTIRMA BULGULARI VE TARTIŞMA

Oyma biçimli karakterlere sahip dairesel metal cisimlerin sınıflandırılması problemine getirilen iki farklı çözüm metaryel ve yöntem bölümünde detaylı bir şekilde açıklanmıştır. Bu yöntemler XOR operatörü kullanılarak şablon eşleştirme işlemi ile sınıflandırma ve ESA modeli kullanılarak karakter tanınması işlemi ile sınıflandırma olarak sunulmuştur. Bu işlemlerin gerçekleştirilmesi için uygulanan algoritmalar ve uygulama işlem süreleri detaylı bir şekilde metaryel ve yöntem bölümlerinde çizelge halinde gösterilmektedir. Çizelgelerin içerikleri örnekleme ürünlerine göre elde edilen işlem süreleri ve doğruluk oranlarına göre tek tek ifade edilmiştir. Bu kısımda ise önerilen (sunulan) yöntemlerin, seçim kriterleri belirlenmesine yönelik çizelgelerin düzenlenmesi ve sonucunda işlem süreleri ve bahsedilecek bazı kriterlere göre en iyi yöntem önerilmiştir.

4.1. Dairesel Metal Cisimlerin Yerini Belirleme Problemi

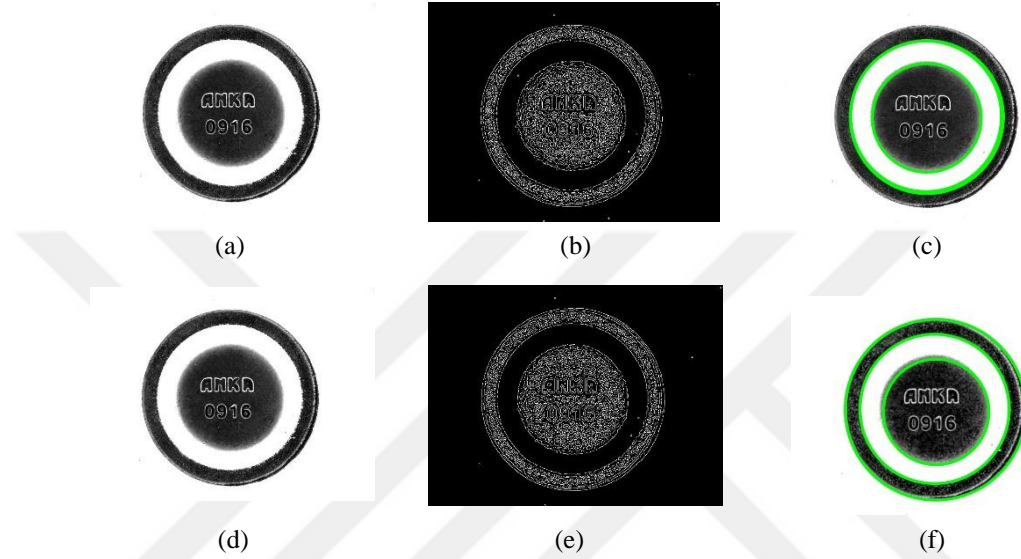
XOR operatörü işlemi uygulanması öncesinde, DRSM çıkarımı için dairesel metal cisimlerin buldukları görüntü içerisindeki yerinin belirlenmesi probleminin çözümü için görüntü içerisindeki dairesel nesne kenarlarının piksel ilişkilerini kullanan Ylajaaski ve Kiryati (1994); Lam ve Yuen (1996) tarafından geliştirilen DHD algoritmasının uygulanmıştır. Yer belirleme problemi öncesinde ise, giriş görüntülerin sahip olduğu gürültülerden temizlemek veya azaltmak, görüntünün yumuşatılması ve görüntü içerisindeki dairesel metal cisimin kenar piksel kontürlerini belirgenleştirmek için Marr ve Hildreth (1980) tarafından bulunan Laplace LoG filtresi uygulanmıştır. Elde edilen filtre görüntüsüne, dairesel metal cisimlerin daire kenarlarının piksel noktalarının tespiti için Canny (1986) tarafından bulunan Canny kenar bulma algoritması uygulanmıştır. Dairesel metal cisimlerin yerini belirleme problemi için uygulanan tüm algoritmalar aşağıda verilmiştir.

4.1.1. Laplacian of Gaussian (LoG) Filtresi

Marr ve Hildreth (1980) tarafından bulunan LoG filtresi kullanılarak, giriş görüntüsü yumuşatılıp görüntü üzerindeki gürültülerin azaltılmaya çalışılmış ve görüntü içerisindeki dairesel metal cisimin kenar kontürlerini keskinleştirmek için uygulanmıştır. LoG piksel değerleri, giriş görüntüsünün piksel değerlerini şu şekilde hesaplamaktadır.

$$LoG(x, y) = -\frac{1}{\pi\sigma^4} \left[1 - \frac{x^2 + y^2}{2\sigma^2} \right] e^{-\frac{x^2 + y^2}{2\sigma^2}} \quad (4.1)$$

Denklem (4.1)' de, (x,y) giriş görüntüsünün pikselleri, σ standart sapma değerini, $LoG(x, y)$ ise çıkış görüntüsü veya filtre uygulanmış görüntünün piksel değerlerini ifade etmektedir. Gaussian Filtresinin σ değeri 4 olarak seçilmiş ve bu oranda giriş görüntüsü yumuşatılmıştır. Görüntüdeki kenar piksel bilgilerinin hesaplanması için LoG filtresi uygulanmıştır. LoG filtresi uygulanmadan elde edilen DHD sonucu ve uygulandıktan sonra elde edilen DHD sonucu Şekil 4.1.' de verilmiştir.

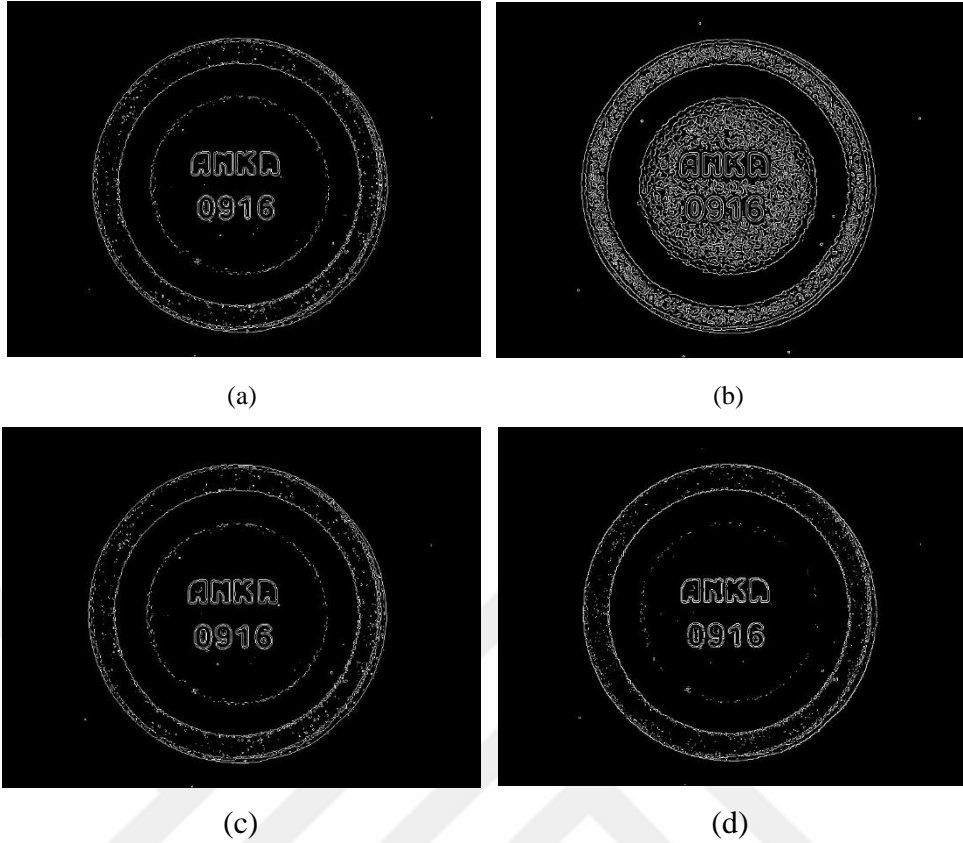


Şekil 4.1. (a)(d) Kamera görsel sonuçları, (b) LoG filtresi uygulanmadan alınan Canny görsel sonucu, (c) LoG filtresi uygulanmadan bulunan daireler, (e) LoG filtre uygulandıktan sonra alınan Canny görsel sonucu, (f) LoG filtresi uygulandıktan sonra bulunan daireler

4.1.2. Kenar Bulma Operatörü

LoG filtresi sonrası elde edilen görüntüye kenar bulma operatörleri uygulanmıştır. Dairesel metal cisimlerin, daire kenarları kontür bilgilerini koruyan ve en iyi sonucu veren kenar bulma operatörü, uygulamadan sonra seçilmiştir. Seçimdeki temel kriter, karakter bölgesi dairesini, orta daireyi ve en dıştaki daire kenar kontür bilgilerini çıkarabilmesi ve yerini belirleme problemi için aynı zamanda dairesel çaplarının bulunması önem taşımaktadır.

Uygulanan kenar bulma operatörleri, Sobel ve Feldman (1968) tarafından bulunan Sobel operatörü, Canny (1986) tarafından bulunan Canny operatörü, Prewitt (1970) tarafından bulunan Prewitt operatörü, Roberts (1963) tarafından bulunan Roberts operatörleridir. Uygulama sonucu elde edilen görüntüler Şekil 4.2.' de gösterilmektedir.



Şekil 4.2. LoG filtresi uygulanan görselin (a) Sobel operatörü sonucu, (b) Canny operatörü sonucu, (c) Prewitt operatörü sonucu, (d) Roberts operatörü sonucu

Şekil 4.2’ de görüldüğü üzere karakter bölgesi dairesinin ve diğer dış dairelerin kenar kontürleri, Canny operatörü uygulanması sonucu, diğer kenar bulma operatörlerine göre daha keskin olup DHD algoritması için en belirgin pikselleri sunmaktadır. Dolayısıyla, örneklenecek görsellere Canny operatörü kenar bulma algoritması uygulanmasına karar verilmiştir. Canny operatörü, kenar piksellerini, şu formüller ile hesaplamaktadır.

$$SNR = \frac{A \left| \int_{-W}^0 f(x) dx \right|}{n_0 \sqrt{\int_{-W}^W f^2(x) dx}} \quad (4.2)$$

$$Localization = \frac{A |f(0)|}{n_0 \sqrt{\int_{-W}^W f^2 dx}} \quad (4.3)$$

$$k = SNR * Localization \quad (4.4)$$

$$k = R_{ij}(\tau) = \int_{-\infty}^{\infty} SNR(x) Localization(x + \tau) dx \quad (4.5)$$

$$R_{ij}(x) = LoG(x, y) e^{\left(-\frac{x^2}{4\sigma^2}\right)} \quad (4.6)$$

Denklem (4.2)’ de, $f(x)$ gürültüyü düzelten ve kenar pixelleri bulmaya yarayan konvolüsyon filtresi, $-W$, W kenar pixelleri, A uygulanan konvolüsyon filtresinin her bir

adımının genliği, n delta fonksiyonunun türev sayısını, SNR kenar bulma operatörünün çıkış sinyal-gürültü oranını (hata oranını) belirtir. SNR değerinin yüksek olması, çıkış fonksiyonunda fazla sinyalin, aynı zamanda az gürültünün olduğunu göstermektedir.

Denklem (4.3)' de ise, *Localization*, bulunan gerçek kenar pixel konumunun, filtrenin yerleştirilen kenar pixel konumu arasındaki mesafenin karşılığını temsil eder. *Localization* değerinin büyük olması, gerçek kenar pixel konumları arasındaki mesafenin küçük olduğunu göstermektedir.

Denklem (4.4)'te SNR ve *Localization* değerlerinin konvolüsyon çarpımını göstermektedir. k , konvolüsyon çarpımını temsil etmektedir. Denklem (4.5)'te ise bu konvolüsyon çarpımı ifade edilirken, i ve j , kartezyen koordinatta piksel değerlerini, R çıkış değerlerini belirtmektedir. Elde edilen k değeri sonucunda, karmaşık yüksek olan değer için Gaussian fonksiyonunun birinci türevine dönülebilmektedir. Denklem (4.6)' da gaussian fonksiyonunun birinci türevi ifade edilirken, işlem sonucu elde edilen gradyen genliği değerlerini içeren kenar haritasına, non-maximum suppression tekniği uygulanarak kenar dışı pikseller ortadan kaldırılmaktadır (Canny, 1986).

Bu şekilde, Canny kenar bulma işlevi, Gaussian filtresinin türevini kullanarak gradyanı hesaplar. Daha sonra, keskin kenarlara bağlı olan zayıf kenarlar olması durumunda, güçlü ve zayıf kenarları tespit etmek için iki eşik kullanır. İki eşik kullanıldığında, Canny operatörünün görüntü içerisindeki gürültüden etkilenmesi diğer yöntemlerden daha az olası ve zayıf kenarları saptama olasılığı daha yüksektir.

Bu bağlamda, zayıf kenarları tespit etmek için eşikleme değeri 8, güçlü kenarlı tespit etmek için eşik değeri 3 seçilmiştir. LoG görüntüsüne uygulanan Canny kenar bulma operatörü sonrası elde edilen görsel Şekil 4.2.(b)' de verilmiştir. Uygulanan konvolüsyon filtresi ise, Şekil 4.3.' de gösterilmektedir.

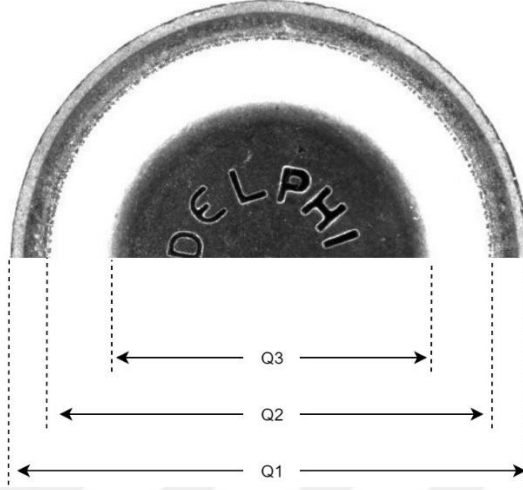
+1	0
0	-1

Şekil 4.3. Canny Operatörü için uygulanan konvolüsyon filtresi

4.1.3. Dairesel Hough Dönüşümü

XOR şablon eşleştirme yaklaşımı için DRSM görselinin çıkarımı için karakter bölgesini, kendi görüntüsüne oranla minimize etmek gerekmektedir. Bu bağlamda daireysel metal cisimlerin öncelikle yerinin belirlenmesi problemi çözülmelidir. Dairesel şekillerinden dolayı, görüntü piksel ilişkileri üzerine çalışmalarıyla ilk defa Hough (1962)

tarafından ortaya atılan cisim kenar piksel ilişkileri, ve teoremin dairesel cisimler için Ylajaaski ve Kiryati (1994); Lam ve Yuen (1996); Smereka ve Duleba (2008) tarafından geliştirilen yöntemlerince DHD algoritmasının uygulanması ile dairesel metal cisimlerin görüntü içerisindeki yerinin tespit edilmesi problemi çözülmüştür.



Şekil 4.4. Dairesel metal cisimlerin daire bölümlerinin gösterimi

Şekil 4.4.'de görüldüğü üzere dairesel metal cisimler, 3 farklı daire ebatı olarak ayrılmıştır. $Q1$ en dış daire çapını, $Q2$ orta daire çapını, $Q3$ ise en iç daire çapını yada karakter bölgesi daire çapını temsil etmektedir. Yerini belirleme problemi çözümlenirken, aynı zamanda daire çaplarının ölçümleride gerçekleştirilmiştir. DHD algoritması, Canny sonucu görüntüsünün piksellerini şu formüller ile hesaplanmaktadır.

$$a(x) = \sum_{r_{min}}^{r_{max}} \sum_{\theta_{min}=0}^{\theta_{max}=360} x - \left(r \times \cos \left(\theta \times \left(\frac{\pi}{180} \right) \right) \right), \quad 0 < x < x_{max} \quad (4.7)$$

$$b(y) = \sum_{r_{min}}^{r_{max}} \sum_{\theta_{min}=0}^{\theta_{max}=360} y - \left(r \times \sin \left(\theta \times \left(\frac{\pi}{180} \right) \right) \right), \quad 0 < y < y_{max} \quad (4.8)$$

$$C(a, b, r) = a(x), b(y), r \quad (\forall x \in a), (\forall y \in b) \quad (4.9)$$

Denklem (4.7) ve (4.8)' de, r_{min} ve r_{max} bulunması muhtemel daire için minimum ve maksimum piksel birimli çap değeri, θ_{min} ve θ_{max} açısal derece, a çemberin x eksen piksel noktası, b çemberin y eksen piksel noktası, r ise çemberin yarı çapını belirtmektedir. Denklem (4.9)' da, $C(a, b, r)$ a ve b işlemleri sürecinde her bir döngü içerisinde oylanan piksel noktalarının sayılması sonucu oluşturduğu çember merkez noktaları ve yarı çapını ifade etmektedir.

$Q1$, $Q2$ ve $Q3$ için farklı r_{min} ve r_{max} değerleri belirlenmiştir. Bu değerler karar verilirken, örnekleme görüntüleri içerisinde en küçük ve en büyük dairesel metal cisimin Eurocambi dijital kumpas cihazı ile alınan daire ölçütlerine göre belirlenmiştir. $Q1$ için

[205, 296], Q_2 için [162, 235], Q_3 için ise [118 173] minimum ve maksimum yarı çap aralıkları belirlenmiştir. Belirlenen parametreler sonrasında, her bir daire için bulunan yarı çap piksel değerleri aşağıdaki formül ile mm cinsine çevrilerek hesaplanmıştır.

$$ratio = \frac{(0.1 \times 33,000)}{49.40} \quad (4.10)$$

$$Q_{1,2,3} = (r_{1,2,3} \times 2) \times ratio \quad (4.11)$$

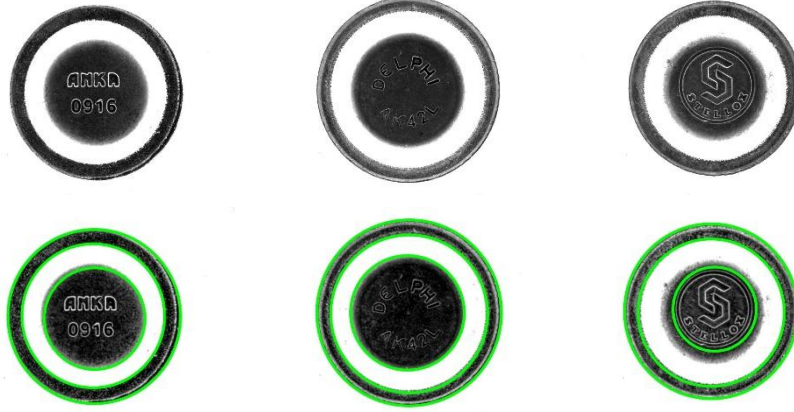
Denklem (4.10)' da *ratio* değeri "ANKA" isimli örnek ürününün görüntü içerisinde kapladığı piksel değeri olan 494 piksel ve 33 mm daire çapı büyüklüğüne göre belirlenmiştir. Denklem (4.11)' de ise DHD algoritması sonrası elde edilen r_1, r_2, r_3 yarı çap değerleri ile işleme sokulan sabit *ratio* değeri sonrası Q_1, Q_2 ve Q_3 çapları bulunmuştur. Böylelikle dairesel metal cisimlerin yerini belirleme problemi çözülmüştür.

Kenarları belirlenmiş görüntüye uygulanan DHD algoritması sonrası bulunan çap değerleri mm cinsinden Çizelge 4.1.'de gösterilmektedir.

Çizelge 4.1. DHD algoritması sonrası tespit edilen daire çapları

Örnek Adı	Q1	Q2	Q3	İşlem Süresi	Doğruluk
ANKA	33 mm	27.26 mm	19.24 mm	3,000 s	% 98.70
DELPHI	35.40 mm	30.06 mm	21.78 mm	3.14 s	% 98.27
STELLOX	33.13 mm	27.92 mm	15.90 mm	2.92 s	% 99,000
4U	-	22.45 mm	22.45 mm	2.44 s	% 96.30
A1016	28.59 mm	24.45 mm	18.44 mm	2.69 s	% 95.42
HEARTH	-	22.45 mm	22.45 mm	2.39 s	% 97.48
WONGA	30.33 mm	30.46 mm	16.30 mm	2.50 s	% 98.58
S11S45	38.21 mm	26.45 mm	18.17 mm	3.52 s	% 99.18
1118	38.48 mm	26.45 mm	16.97 mm	3.72 s	% 98.73
METZGER	38.21 mm	26.72 mm	18.70 mm	3.70 s	% 99.10
APLUS	29.26 mm	29.56 mm	18.57 mm	2.89 s	% 98.59

Çizelge 4.1.'de verilen daire çapları milimetre cinsindedir ve diğer yöntemler ile karşılaştırılması için mevcut ortalama işlem süreleri ve Eurocambi dijital elektronik kumpas cihazı ile alınan değerlere oranla doğruluk yüzdeleri belirtilmiştir. Q_1 dış çapı, Q_2 orta çapı, Q_3 ise iç çapı temsil etmektedir. '-' ifadesi tespit edilemediğini, 4U ve HEARTH örneklerinin Q_3 sonuçları ve APLUS Q_1 sonucunun yanlış tespit edildiği gösterilmektedir. Tespit edilen daire konturları giriş görüntüsü üzerine çizilerek Şekil 4.5.'de verilmiştir.



Şekil 4.5. DHD algoritması sonuçları, tespit edilen daireler işaretlenerek gösterilmektedir

4.2. Daugman's Rubber Sheet Modeli Çıkarımı

Karakter bölgesi dairesi veya tespit edilen $Q3$ isimli en iç daire (en küçük daire) bulunduğu görüntüden kesilerek alınmıştır. Bundan sonraki işlem görüntünün normalize edilerek özellik çıkarımının gerçekleştirilmesi için hazır hale getirilmesidir. Normalize işlemi için Daugman (1997) tarafından geliştirilen bir teknik ile kesilen görüntü polar düzleminden kartezyen düzlemine aktaran DRSM yöntemi kullanılmıştır. DRSM yöntemi, dairesel bir hizada bulunan karakterlerin düz bir hizaya getirilmesi, normalize edilmesi için bu yaklaşım kullanılmıştır. Daugman nesne sınıflandırma problemi içerisinde insan kimliği tespiti için insan gözünü kullanarak şablon eşleştirme yöntemi geliştirmiştir. DRSM görüntüsü, şu formüller ile hesaplanmaktadır.

$$I(x, y) \rightarrow I(x(r, \theta), y(r, \theta)) \rightarrow I(r, \theta) \quad (4.12)$$

$$x(r, \theta) = (1 - r)x_p(\theta) + rx_s(\theta) \quad (4.13)$$

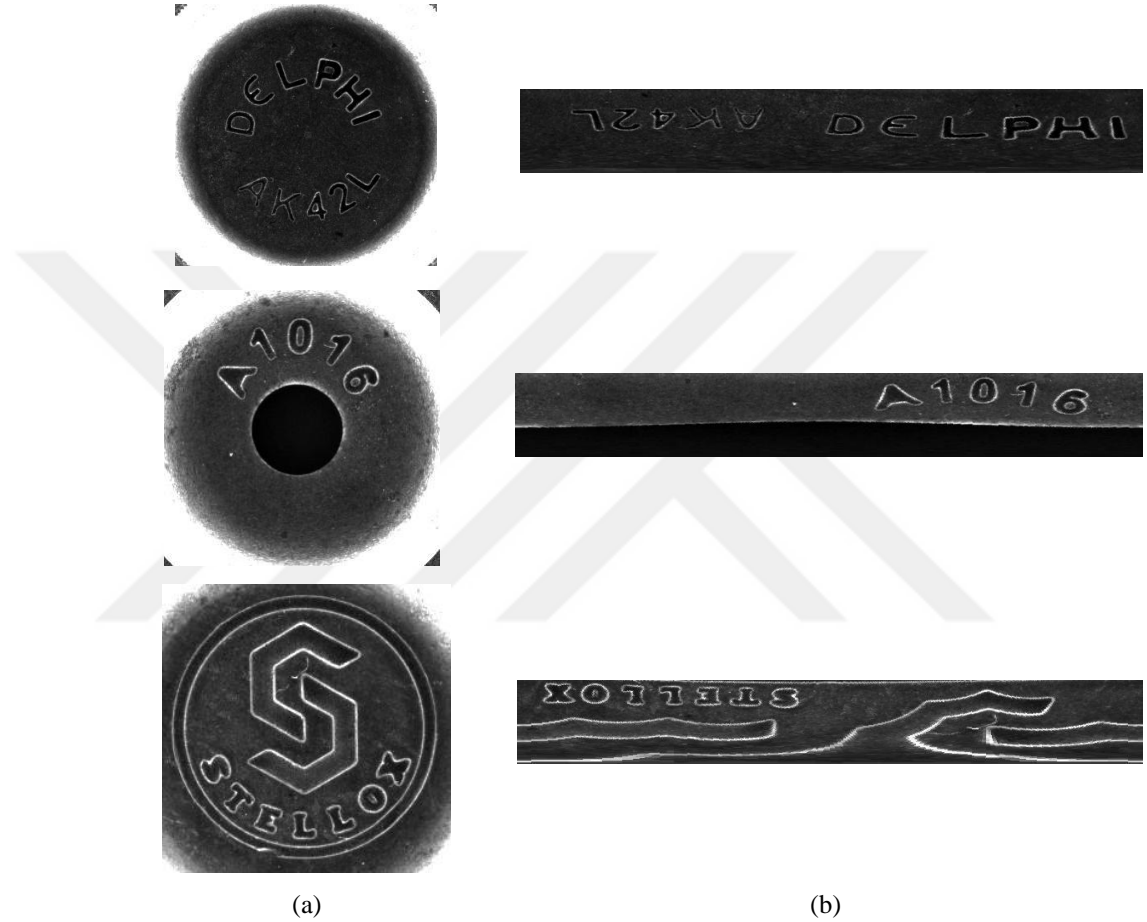
$$y(r, \theta) = (1 - r)y_p(\theta) + ry_s(\theta) \quad (4.14)$$

$$DRS(x, y) = I(x(r, \theta), y(r, \theta)) \quad (4.15)$$

Denklem (4.12) gösteriminde, $I(x, y)$ kartezyen koordinat düzleminde bulunan görseli, $x(r, \theta)$ ve $y(r, \theta)$ ise polar koordinat düzlemini belirtmektedir. Kartezyen koordinat düzleminde bulunan dairelerin r ve θ şeklinde gösterimleri, ve polar koordinat düzlemi içinde $I(r, \theta)$ gösterimlerini belirtmektedir.

Denklem (4.13) ve (4.14)' te, $x_p(\theta)$ ve $y_p(\theta)$ karakter bölgesinin dışarıda kalmayacak bir şekilde görüntünün açılması (image warping) için belirlenen karakterlerin dip kenar piksel noktaları, $x_s(\theta)$ ve $y_s(\theta)$ ise kesip çıkarılan en küçük dairenin dış çaplarının piksel noktalarını belirtmektedir. Denklem (4.15)' te elde edilen normalize koordinat düzleminin piksel noktalarını $DRS(x, y)$ ifade etmektedir.

Denklemleri verilen DRSM çıkarımı için, $x_p(\theta)$ ve $y_p(\theta)$ piksel noktaları için, en küçük dairenin yarıçap değerinden 30 piksel daha az olarak verilmiştir. Çünkü, tam olarak karakter bölgesi, en küçük dairenin biraz daha içerisinde bulunmaktadır (Şekil 4.5.). $x_s(\theta)$ ve $y_s(\theta)$ piksel noktaları ise, en küçük dairenin merkez noktaları olarak belirlenmiştir. Belirlenen değerler sonrasında elde edilen DRSM görüntüsü Şekil 4.6. (b)'de gösterilmektedir.



Şekil 4.6. (a) DHD sonrası en iç daire alanının kesilip alınmış görüntüleri, (b) DRSM yöntemi ile normalize edilmiş (kartezyen) görüntüler

En iç daire bölgesi alınan kare biçimli görüntü DRSM işlemi sonrasında dikdörtgen biçimli olmuştur. Örneğin 'DELPHI' örnek isimli görüntü için 327x327 çözünürlüğünden normalize edildikten sonra 99x750 çözünürlüğe uzatılmıştır. Polar görüntünün, kartezyen görüntüsü için yeniden belirlenen piksel noktaları DRS ile oluşturulan normalize edilmiş görüntüler, bir sonraki aşama olan özellik çıkarımı için hazır hale getirilmiştir.

4.3. Ortalama Mutlak Sapma (OMS) ile Özellik Çıkarımı

DRSM çıkarımı gerçekleştirilmiş görüntüler veri tabanına şablon eşleştirme için kaydedilmeden önce, görüntülerin özellikleri çıkarımları gerçekleştirilmiştir. Bu işlem için, Turner (1986) tarafından ortaya konan doku analizi, Havlicek ve ark. (1997) yaptıkları çalışmada belirtilen belirli bant-geçiş kanalları ve uzaysal frekanslar gibi parametreler veren Gabor (1946) tarafından bulunan Gabor Filtre (GF) uygulanmıştır. GF şu formüller ile DRSM görüntü piksel değerlerine uygulanmaktadır.

$$x' = x \cos \theta + y \sin \theta \quad (4.16)$$

$$y' = y \cos \theta - x \sin \theta \quad (4.17)$$

$$G(x, y; \theta, f) = \exp \left\{ -\frac{1}{2} \left[\frac{x'^2}{\delta_x^2} + \frac{y'^2}{\delta_y^2} \right] \right\} \cos(2\pi f x') \quad (4.18)$$

Denklem (4.16) ve (4.17)' de, x' ve y' giriş görüntü piksel değerlerinin radyan değerlerini, θ ise filtrenin uygulanacağı açı değerini ifade etmektedir. Denklem (4.18)' de, f dalga uzunluğunu (wavelength - frekans - cycles/degree), δ değerleri kontrast sapma, $G(x, y)$ giriş görüntüsünün GF uygulandıktan sonraki magnitude görüntüsü, $G(\theta, f)$ ise giriş görüntüsünün GF uygulandıktan sonraki faz görüntüsünü belirtmektedir. GF uygulanırken $f = 2, 4, 8, 16, 32$ olarak, $\theta = 0^\circ, 45^\circ, 90^\circ, 135^\circ$ olarak belirlenmiştir. Uygulamada en iyi sonuçları edinilen, frekans ve θ sabitleri için (Tan, 1995; Lee, 1996; Clausi ve Jernigan, 2000) çalışmalarından faydalanılmıştır. Özellik çıkarımı için desen, şablon faktörlerinin öne çıktığı GF' nin faz görüntüleri kullanılmıştır (Oppenheim ve ark., 1983; Scott T. Acton ve ark., 2001; Tisse ve ark., 2002). Elde edilen faz görüntüleri Şekil 4.7.'de gösterilmektedir. Faz görüntüleri, her bir örnek için 5 farklı f frekans ve 4 farklı θ açı değerleriyle 20 adet, 11 örnekleme görüntüsü için toplamda 220 adet oluşturulmuştur. GF sonrası elde edilen faz görüntülerin renk yoğunluk aralığı -3.14 ile 3.14 olarak tespit edilmiş olup, şablon eşleştirme için kullanılması gereken "0" ve "1" ikili değerlere dönüştürülürken, değer kaybının minimuma indirgenmesi için Ma ve ark. (2002) tarafından gerçekleştirilen bir çalışmadan esinlenerek faz görüntülere OMS algoritması uygulanmıştır. OMS algoritması GF phase görüntüsünün piksellerini şu formüller ile hesaplamaktadır.

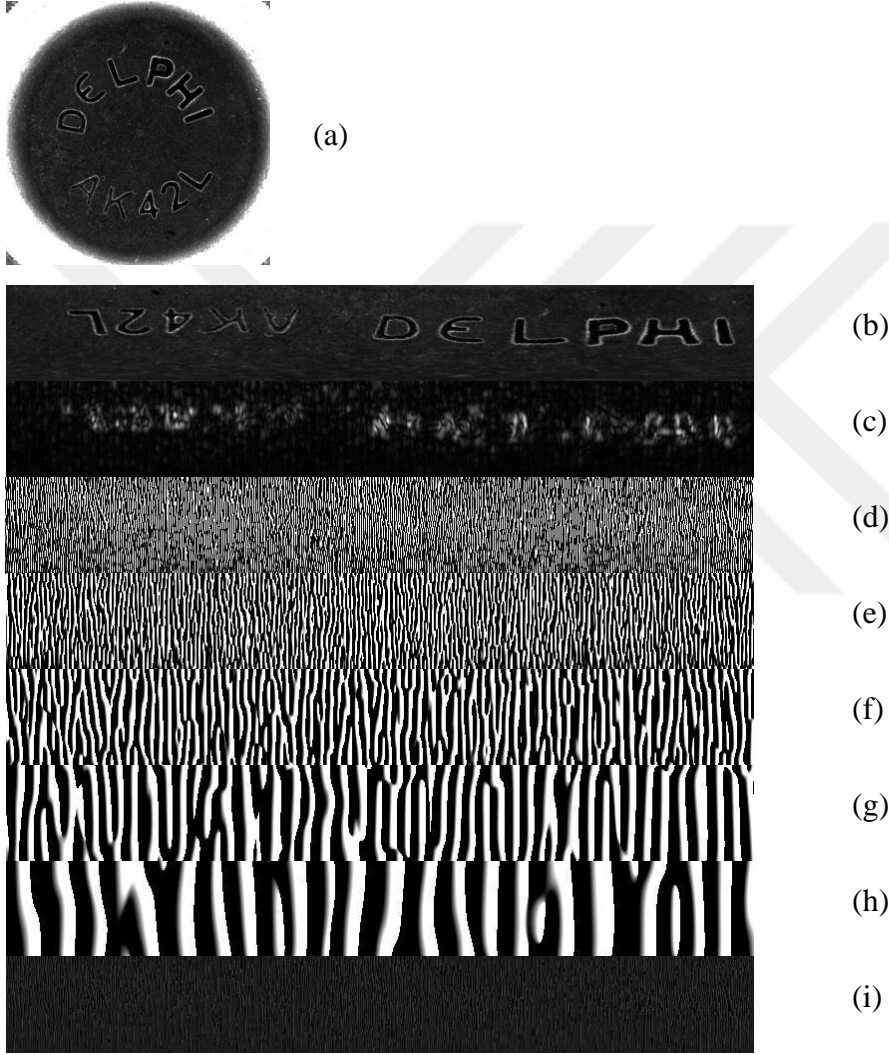
$$OMS(x, y) = \frac{1}{N} \sum_{\theta, f=1}^N |G(\theta, f) - \text{mean}(G(\theta, f))| \quad (4.19)$$

$$\sigma = 1.253 \times OMS(x, y) \quad (4.20)$$

$$OMS'(x, y) = G(\theta, f) \times \sigma \quad (4.21)$$

Denklem (4.19)' da, x ve y phase görüntüsünün pixel değerlerini, N piksel değerlerinin toplamını, $G(\theta, f)$ ise elde edilen GF' ninin faz görüntüsünü belirtmektedir.

Denklem (4.20)' de, σ OMS(x, y) görüntüsünü normalleştirmek için kullanılan bir sabittir. Denklem (4.21)' de $OMS'(x, y)$ ise elde edilen çıkış görüntüsünü ifade etmektedir. Uygulanan OMS algoritma sonrası, elde edilen faz görüntülerin pixel değerleri, "0" ın altında ise 0, "0" ın üstünde ise 1 olacak şekilde eşiklenerek görüntünün ikili dönüşümü gerçekleştirilmiştir. OMS algoritmasının uygulanması sonucu elde edilen görüntü Şekil 4.7. (i)'de verilmiştir.



Şekil 4.7. (a) DHD sonrası en iç daire alanının kesilip alınmış görüntüsü, (b) En iç daire alanının DRSM çıkarımı, (c) GF $\theta = 0$, $f = 2$ uygulanmış magnitudo görüntüsü, (d) GF $\theta = 0$, $f = 2$ uygulanmış faz görüntüsü, (e) GF $\theta = 0$, $f = 4$ uygulanmış faz görüntüsü, (f) GF $\theta = 0$, $f = 8$ uygulanmış faz görüntüsü, (g) GF $\theta = 0$, $f = 16$ uygulanmış faz görüntüsü, (h) GF $\theta = 0$, $f = 32$ uygulanmış faz görüntüsü, (i) GF $\theta = 0$, $f = 2$ uygulanmış faz görüntüsünün OMS algoritması sonrası özellik çıkarımı görüntüsü

Şablon eşleştirme için işlem süresinin kısaltılması maksatıyla, görüntü bilgisi 99x750 - 74250 piksel' den 10x75 - 750 piksele düşürülmüştür. Tekrar boyutlandırılan görseller tek bir hizada dizi haline getirilerek 1x750 çözünürlüğünde veri tabanına, daha

sonra eşleştirilmek üzere eklenmiştir. Veri tabanına ayrıca ilgili örnek ismi ve index numarası eklenmiştir.

4.4. XOR Şablon Eşleştirme

Veri tabanına kaydedilen özellik çıkarımı gerçekleştirilmiş görüntüler, her bir yeni giriş görüntüsü, eşleştirilmek üzere sunulduğu anda aynı işlemler gerçekleştirilerek, ikili görüntüleri XOR operatörü ile işleme sokulur. XOR lojik operatörü, aynı ikili değerlerin işlemi sonucu verdiği "0" yanıtı ile şablon eşleştirme amacıyla kullanılan oldukça kullanışlı ve hızlı bir yöntemdir (Zhang ve ark., 2010; Bala ve Tajinder, 2016). XOR operatörü işlemi denklem (4.22)' de hesaplanırken, eşleşme oranı veren α ise denklem (4.23)' deki formül ile hesaplanmaktadır.

$$S = \sum_{i=1}^N A_i \oplus B_i \quad (4.22)$$

$$\alpha = \frac{1}{N} \times S(0) \quad (4.23)$$

Denklem (4.22)' de, i eşleştirilecek görüntü ve veritabanındaki görüntünün piksel indeksini, N toplam index sayısını, A eşleştirilecek görüntüyü, B veritabanındaki görüntüyü, S , XOR operatörü işlemi sonrasında elde edilen görüntüyü ifade etmektedir. Denklem (4.23)' te ise, $S(0)$ görüntünün "0" ikili değerli piksellerin toplamını, α ise eşleşme oranını belirtmektedir. Örnekleme görüntülerinin uygulandığı XOR işlemi sonrasında elde edilen eşleşme oranları, DRSM çıkarımı, GF uygulaması ve XOR operatörü işlemleri için işlem süreleri Çizelge 4.2.' de gösterilmektedir.

Çizelge 4.2. XOR operatörü işlemi sonrası eşleşme oranları

ÖRNEK ADI	DRSM süresi	GF süresi	XOR süresi	Eşleşme Oranı
ANKA	27 ms	4 ms	37 ms	0.932687
DELPHI	34 ms	4 ms	38 ms	0.894114
STELLOX	22 ms	3 ms	39 ms	0.960512
4U	39 ms	4 ms	39 ms	0.963340
A1016	27 ms	4 ms	39 ms	0.876377
HEARTH	35 ms	4 ms	39 ms	0.965684
WONGA	22 ms	3 ms	37 ms	0.932081
S11S45	24 ms	4 ms	38 ms	0.962653
1118	23 ms	3 ms	38 ms	0.955879
METZGER	27 ms	4 ms	38 ms	0.931286
APLUS	27 ms	3 ms	38 ms	0.963246

Çizelge 4.2.’ de XOR şablon eşleştirme yöntemi için uygulanan algoritalarda, örnek ürünlerin işlem süreleri ve eşleşme oranları belirtilmektedir. DRSM süresi, DSRM için yapılan işlem süresini, GF süresi, özellik çıkarımı için uygulanan Gabor Filtresi için yapılan işlem süresini, XOR süresi, şablon eşleştirme için yapılan işlem süresini, Eşleşme Oranı ise XOR işlemi sonrasında elde edilen eşleşme oranları belirtmektedir.

4.5. Maksimum Kararlı Ekstrimal Bölgeler (MSER) Uygulanması

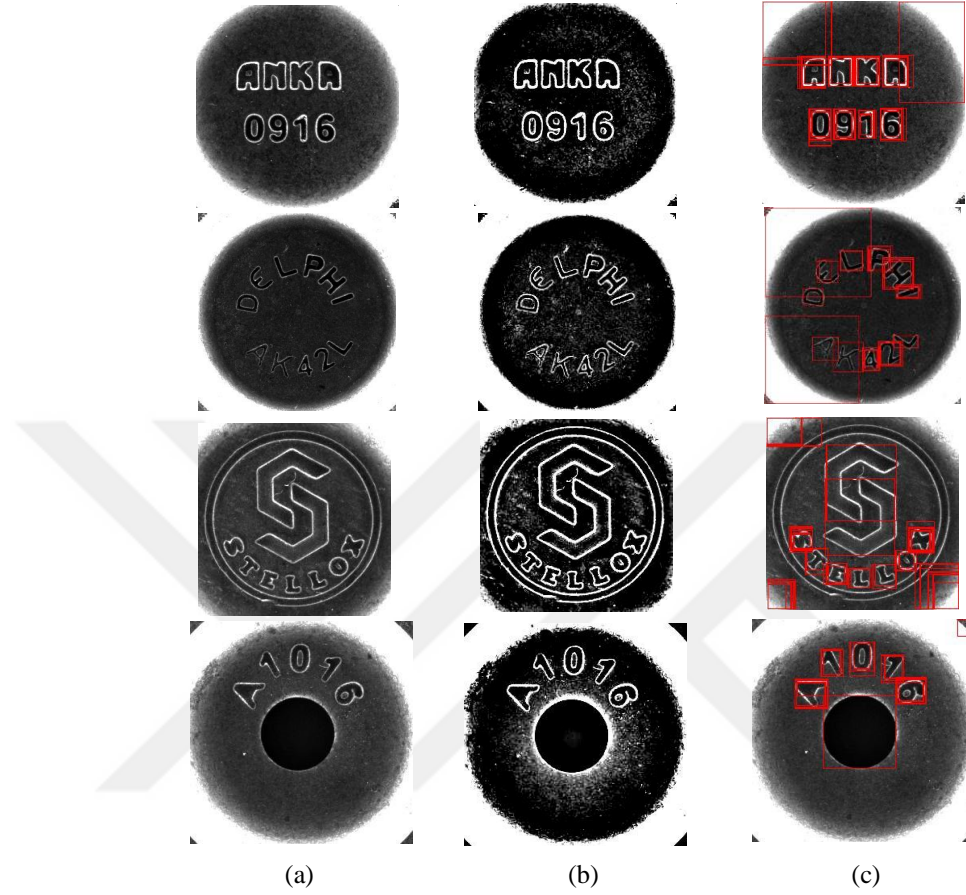
Karakter bölgesi tespiti ön işleminden sonra, karakter bölgesinin daha iyi bir şekilde tespit edilebilmesi, karakterlerin görüntü içerisinde keskinleştirilmesi için ön karakter bölgesi görüntüsüne, kontrast genişletme işlemi için keskinleştirme filtresi (SF) uygulanmıştır (Erdinc Kocer ve Kursat Cevik, 2011). SF uygulanmış görüntünün sonucu için Şekil 4.8 (b)’de verilmiştir. SF uygulanması sonrasında, elde edilen görüntüde karakter kenar piksellerin kontrastı artırılmış olup, karakter bölgesi tespiti ve segmentasyonu için kullanışlılığı artırılmıştır. Karakter bölgesi çıkarımı ön işlemi işlemi DHD, karakter kontür piksellerinin keskinleştirilmesi için SF uygulandıktan sonra, karakter segmentasyonu, diğer bir deyiş ile karakterlerin tek tek ayırt edilebilmesi için MSER algoritması kullanılmıştır. Temelleri Lindeberg (1993) tarafından atılan Blob Analizi (Blob Coloring) algoritması, görüntü içerisindeki, görüntünün arka planından farklı nesnelerin (kararlı nesnelerin) tespit edilmesini veya tanınmasını sağlamaktadır. Blob analizi yöntemlerinden olan ve Matas ve ark. (2004) tarafından geliştirilen MSER algoritması, SF işlemi sonucunda elde edilen görüntüye uygulanarak, karakterlerin segmentasyonu gerçekleştirilmiştir (Salembier ve Garrido, 2000; Mikolajczyk ve ark., 2005; Donoser ve Bischof, 2006). MSER algoritması, uygulanan görselin piksellerinin alanlarını şu formül ile hesaplamaktadır.

$$\psi(R_i^g) = \frac{|R_i^{g-\Delta} - R_i^{g+\Delta}|}{|R_i^g|}, \quad (\Delta \in \mathcal{S}) \quad (4.24)$$

Denklem (4.24)’ de, i [0-255] arasındaki her bir eşik değerine göre görüntü içerisinde elde edilen nesnelerin sayısı, g [0-255] arasındaki eşik değerleri, R_i^g elde edilen nesnelerin alanları, Δ eşik örnekleme aralığını kontrol etmek için değişken bir parametre, ψ ise karar değeri olarak ifade edilmektedir.

MSER algoritması için SF görüntüsü öncelikle, en iyi eşik değeri bulunan “63-255” arası ile ikili görüntüye dönüştürülmüştür. Δ değeri ise 4 olarak belirlenmiştir. Daha sonra MSER algoritması, ikili piksel arasında MSER karar değeri ile lokal minimum bölgeleri (karakter dışı) bölgeleri engelleyerek, muhtemel karakter kenar kontürlerini

işaretlemiştir. İşaretlenen bu piksel noktaları ile karakter bölgesi sınırlayıcı kutu (bounding box) içerisine alınmıştır. MSER algoritması uygulanmış görüntülerin sonuçları Şekil 4.8. (c)' de gösterilmektedir.



Şekil 4.8. (a) DHD sonrası en iç daire alanının kesilip alınmış görüntüsü, (b) SF uygulaması sonrası karakterlerin keskinleştirilmiş görüntüsü, (c) MSER algoritması uygulanmış, karakter segmentasyonu gerçekleştirilmiş görüntüler

4.6. Kontür Genişliği Dönüşümü (SWT) Gerçekleştirilmesi

Karakter segmentasyonu gerçekleştirilmiş görüntülerde, bazı karakter dışı bölgelerde sınırlayıcı, kutu içerisine alınarak işaretlenmiştir. MSER algoritması, karakter alanlarını tespit ederken, aynı zamanda muhtemel karakter dışı alanları da tespit etmiştir. Bu durumdan kurtulmak için, Epshtein ve ark. (2010) tarafından geliştirilen, karakter dışı olan bölgeleri engelleyebilen ve görüntü dışına çıkarabilen bir algoritma olan SWT algoritması uygulanmıştır. SWT, karakterlerin kontür genişliği bilgisini kullanarak, görüntü içerisinde bu bilgidan fazla veya az olan kontür genişliği bilgilerini elemektedir. SWT algoritması, uygulanan görselin piksellerini şu formüller ile hesaplamaktadır.

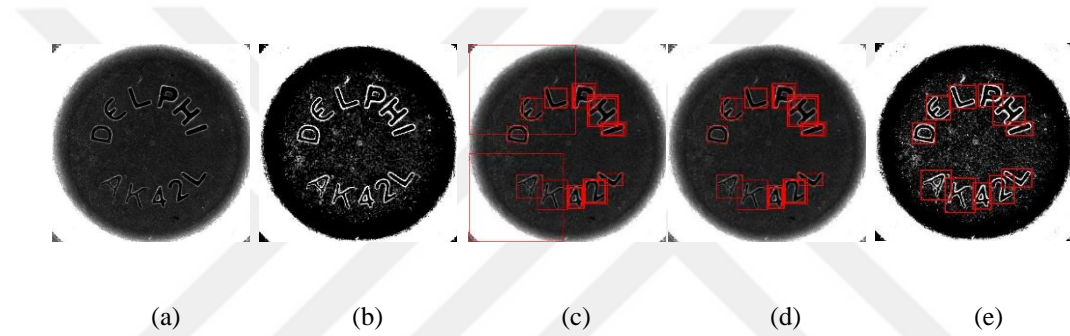
$$r = p + \sum_{p=1}^N nd_p \quad (4.25)$$

$$d_q = -d_p \pm \frac{\pi}{6} \quad (4.26)$$

$$s = p - q \quad (4.27)$$

Denklem (4.25), (4.26) ve (4.27)' de p karakterin iç kontür pikselini, N , p' nin toplam piksel sayısını, n iterasyon sayısını, d_q ve d_p piksellerin gradyan yönlerini, q bulunan karakterin dış kontür pikselini, s ise bulunan p ve q piksellerin arasını, yani piksel genişliğini ifade etmektedir.

SWT algoritmasının uygulanması sonrasında, elde edilen karakter piksel kontürleri ile karakter olmayan bölgeler elenmiş ve karakter bölgeleri yeniden işaretlenmiştir. SWT algoritması uygulanmış görüntülerin sonucu Şekil 4.9. (d)' de gösterilmektedir.



Şekil 4.9. (a) DHD sonrası en iç daire alanının kesilip alınmış görüntüsü, (b) SF uygulaması sonrası karakterlerin keskinleştirilmiş görüntüsü, (c) MSER algoritması sonrası, karakter segmentasyonu gerçekleştirilmiş görüntü, (d) SWT algoritması sonrası, karakter dışı bölgelerin elenmesi, (e) Normalize edilen karakter kontür bölgesi sonuçları

SWT sonrasında, işaretlenen karakter bölgelerinde, bazı karakterlerin bir kaç kez işaretlenmesi problemi görülmüştür. Bu problemin çözümü için bir dizi matris işlemleri gerçekleştirilmiştir. Bu matrix işlemleri sırasıyla, aynı koordinata sahip olanların elenmesi, x , y koordinat toplamalarının $x+1$, $y+1$ koordinat toplamalarının farkı eğer 10 pikselden küçük ise elenmesi şeklindedir. Bu koordinatlar giriş görüntüsünün koordinatları olmayıp, tespit edilen karakterlerin sınır işaret koordinatlarıdır. Uygulanan matris işlemleri aşağıdaki formüller ile hesaplanmaktadır.

$$bbox(x, y) = \begin{cases} \sum_{x,y=1}^N x, y = \sum_{x,y=2}^{N-1} x, y & , \text{eşitlik ise } 0 \\ \sum_{x,y=1}^N |(x + y) - (x + 1, y + 1)| \leq 10 & , \text{eşitlik ise } 0 \end{cases} \quad (4.28)$$

Denklem (4.28)' de, x ve y tespit edilen karakter sınırı işaret koordinatları, N koordinatlar toplamı, "10" tekrar eden karakter kontür noktaları için belirlenmiş bir sabit değerdir. Karakter tespiti sınır işaret koordinatlarını ise $bbox(x, y)$ temsil etmektedir. Denklem (4.28)' deki eşitlik koşullarının sağlanması ile ilgili sınır işaret koordinatları elenmiştir. Gerçekleştirilen bu işlem, karakter bölgelerinin normalize edilmesi şeklinde

ifade edilebilir. SWT sonrası karakter kontür piksel koordinatları bu formül ile normalize edilerek, karakter bölgeleri yeniden işaretlenmiştir. Normalize edilen karakter kontür piksel koordinatları Şekil 4.9. (e)' de gösterilmektedir. Karakter bölgesi tespiti problemi ve karakter segmentasyonu problemi en son elde edilen normalize koordinat değerleri ile çözülerek, karakterler buldukları görselden kesilip karakter tanınması problemi için hazır hale getirilmişlerdir.

4.7. Optik Karakter Tanınmasının (OCR) Gerçekleştirilmesi

Karakterlerin segmentasyonu sonucu elde edilen karakter bölgeleri karakter tanınması için hazır getirilmiştir. Karakter tanınması ve sonrasında gerçekleştirilecek dairesel metal cisimlerin sınıflandırma problemi için OCR yöntemi kullanılmıştır. İlk tanımı D'Albe (1920) tarafından konulan ve icat edilen Optophone, daha sonra günümüz karakter tanınması için kullanılan OCR adını, Ray Kurzweil tarafından 1976 yılında icat ettiği OCR makinesinden almıştır (Schantz, 1982). OCR işleminin gerçekleştirilmesi için ise, yapay zekanın derin öğrenme sınıflarından olan Evrişimli Sinir Ağı (ESA) algoritması kullanılmıştır (LeCun ve ark., 1998).

ESA algoritmaları karakter tanınması için sıklıkla kullanılan yöntemlerden olup, başarısı ve elde ettiği tahmin doğruluk oranları ile hassasiyetini ispatlamaktadır (Bell ve Sejnowski, 1997; Connell ve Jain, 2001; Hinton ve ark., 2006; Keysers ve ark., 2007; Lee ve ark., 2009; Radzi ve Khalil-Hani, 2011; Krizhevsky ve ark., 2012; LeCun ve ark., 2015; Jailin Reshma ve ark., 2016; Driss ve ark., 2017; Zhou ve ark., 2017).

ESA modeli, görüntü içerisinde karakter boyutları ve karakter pozisyonları temel alınarak tasarlanmıştır. Tasarlanan ESA modeli Şekil 4.10.' da gösterilmektedir. ESA modelinin katmanları maddeler halinde aşağıda sıralanmaktadır.

- Giriş katmanı: Karakter görüntüleri farklı çözünürlüğe sahip olup, modelin eğitilmesi için kullanılan EMNIST eğitim veri setlerinin çözünürlüğü olan, 28x28 çözünürlüğüne görüntüler boyutlandırılarak gri seviye olarak kaydedilmiştir. Dolayısıyla ESA'nın giriş katmanı 28x28x1 şeklinde belirlenmiştir.
- 1. Convolution katmanı: 3x3 konvolüsyon filtresine sahip olan bu katman, giriş görüntüsüne 32 adet filtre uygulamaktadır. Giriş kenar dolgusu (padding) değeri sabit olarak belirlenmiş olup, katmanın çıkış görüntüsü çözünürlüğü sabit tutulmuştur. Böylelikle katmanın aktivasyon haritası (öznitelik haritası) 28x28x32 şeklinde oluşturulmuştur. Bu şekilde katmanın parametre sayısı (ağırlığı) ise $w = ((3 \times 3 \times 1) + 1) \times 32 = 320$ dir. Filtre başına "+1" önyargı (bias) ile $1 \times 1 \times 32 = 32$ adet

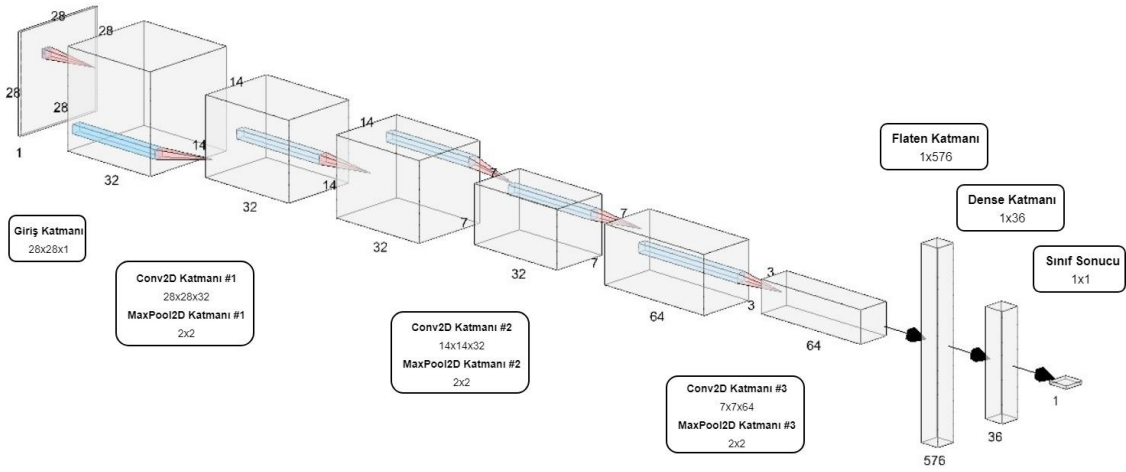
öğrenilebilir nöron bağlantısı oluşturulmuştur. Ayrıca, katmanın aktivasyon fonksiyonu Qian ve ark. (2018) tarafından geliştirilen Relu fonksiyonu seçilmiştir.

- Batch Normalization katmanları: Ioffe ve Szegedy (2015) tarafından geliştirilen, her katmanın aktivasyon fonksiyonunu normalize eden, ortalama aktivasyon değerini $0'$ a, aktivasyon standart sapma değerini $1'$ e yakınsayan işlemdir.
- 1. Max Pooling katmanı: Konvolüsyon filtresi sonrasında maksimum değerleri alır. 2×2 şeklinde belirlenerek, katmanın aktivasyon haritası $14 \times 14 \times 32$ şeklinde oluşturulmuştur. Giriş görüntüsünde filtrenin adım boyutu (stride) değeri 2 seçilmiştir.
- 2. Convolution katmanı: 3×3 konvolüsyon filtresine sahip olan bu katman, giriş görüntüsüne 32 adet filtre uygulamaktadır. Giriş kenar dolgusu (padding) değeri sabit olarak belirlenmiş olup, katmanın çıkış görüntüsü çözünürlüğü sabit tutulmuştur. Böylelikle katmanın aktivasyon haritası (öznitelik haritası) $14 \times 14 \times 32$ şeklinde oluşturulmuştur. Bu şekilde katmanın parametre sayısı (ağırlığı) ise $w = ((3 \times 3 \times 32) + 1) \times 32 = 9248$ dir. Filtre başına “+1” önyargı (bias) ile $1 \times 1 \times 32 = 32$ adet öğrenilebilir nöron bağlantısı oluşturulmuştur. Aktivasyon fonksiyonu olarak Relu fonksiyonu kullanılmıştır.
- 2. Max Pooling katmanı: Ortaklanacak bölgenin boyutu 2×2 şeklinde belirlenerek, katmanın aktivasyon haritası $7 \times 7 \times 32$ şeklinde oluşturulmuştur. Giriş görüntüsünde filtrenin adım boyutu (stride) değeri 2 seçilmiştir.
- 3. Convolution katmanı: 3×3 konvolüsyon filtresine sahip olan bu katman, giriş görüntüsüne 64 adet filtre uygulamaktadır. Giriş kenar dolgusu (padding) değeri sabit olarak belirlenmiş olup, katmanın çıkış görüntüsü çözünürlüğü sabit tutulmuştur. Böylelikle katmanın aktivasyon haritası (öznitelik haritası) $7 \times 7 \times 64$ şeklinde oluşturulmuştur. Bu şekilde katmanın parametre sayısı (ağırlığı) ise $w = ((3 \times 3 \times 32) + 1) \times 64 = 18496$ dir. Filtre başına “+1” önyargı (bias) ile $1 \times 1 \times 64 = 64$ adet öğrenilebilir nöron bağlantısı oluşturulmuştur. Aktivasyon fonksiyonu olarak Relu fonksiyonu kullanılmıştır.
- 3. Max Pooling katmanı: Ortaklanacak bölgenin boyutu 2×2 şeklinde belirlenerek, katmanın aktivasyon haritası $3 \times 3 \times 64$ şeklinde oluşturulmuştur. Giriş görüntüsünde filtrenin adım boyutu (stride) değeri 2 seçilmiştir.

- Fully Connected katmanı: En son elde edilen aktivasyon haritasının 64 nöronu, bu katmanda belirlenen 36 nörona bağlanmıştır. Katmanın toplam parametre sayısı $w = (64 + 1) \times 36 = 2340$ ' dır.
- Softmax katmanı: Bu katman, bir önceki katmanda oluşan skorları, olasılık vektörü şeklinde sunan bir fonksiyondur (Bishop, 2006). Softmax Denklem (4.29)'daki gibi hesaplanmaktadır.

$$p_i = \frac{e^{x_i}}{\sum_{j=1}^n e^{x_j}} , \quad 0 \leq p \leq 1 \quad (4.29)$$

Denklem (4.29)' da, i ve j giriş değerleri indeksini, x giriş değerlerini, n toplam giriş değerlerini, p ise olasılık değerlerini belirtmektedir.

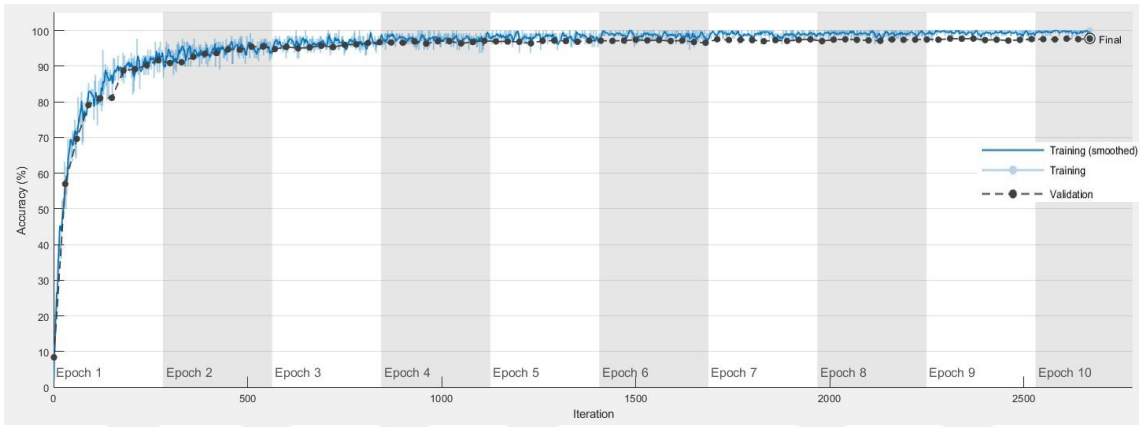


Şekil 4.10. Tasarlanan ESA modelinin isotropik gösterimi

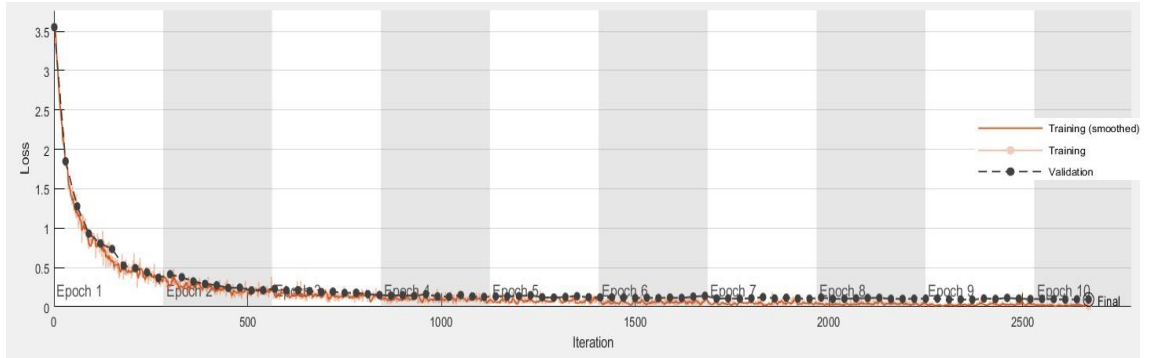
Tasarlanan modelin eğitim süreci için uygulanan bir çok parametre vardır. Bu parametreler opsiyonel parametreler olup, ilgili değerler belirlenmiştir. Optimizasyon fonksiyonu olarak, Kingma ve Ba (2015) tarafından geliştirilen Adam fonksiyonu kullanılmıştır. Öğrenme oranı parametresi 0.01, momentum oranı ise 0.9 olarak belirlenmiştir. Eğitimin dönem sayısı (epoch) parametresi 10, iterasyon sayısı 2810 olarak belirlenmiştir. Yapılan parametre ayarları sonrasında tasarlanan model eğitime başlatılmıştır. Modelin eğitimi için iki farklı veri seti kullanılmıştır. Normalize edilen karakter kontür bölgesi görüntülerin bir kısmı, ikili tersleme, x ve y eksenlerinde ters çevirme, döndürme, gürültü ekleme, histogram eşitleme, adaptif histogram eşitleme, görüntü yoğunluk değerlerini değiştirme işlemleri uygulanarak çoğaltılmış, bir diğer kısmında ise Cohen ve ark. (2017) tarafından oluşturulan EMNIST veri setleri kullanılmıştır. Sınıflandırma ismi (etiket ismi) 36 adet olan, her biri için 1000 adet olmak

üzere 36000 adet eğitim görüntüsü, her biri 250 adet olmak üzere 9000 adet doğrulama görüntüsü kullanılmıştır.

Eğitim süreci için ayarlanan giriş görüntüsü 28x28 çözünürlüğünde gri seviye görüntüler olup, tasarlanan modele göre her bir katmanın uygulanması sonucu, nöronlara 3x3x64 çözünürlük bilgisi ulaştırılmıştır. Nöronların sonucunda 36 adet sınıf bilgisi tahmin oranları elde edilerek, en yüksek orana sahip olan ifade tahmin edilen veya karakteri tanınan bilgisini vermektedir. Eğitim sonucunda minimum doğrulama kaybı 0.088 oranında, doğruluk yüzdesi %97.76, doğrulanmış doğruluk yüzdesi ise %97.66 olarak elde edilmiştir. Doğruluk yüzdeleri ve kayıp oranlarını içeren eğitim süreci grafikleri Şekil 4.11.'de gösterilmektedir.



(a)



(b)

Şekil 4.11. Eğitim sürecinde (a) artan doğruluk yüzdesi ve doğrulama doğruluk yüzdesi değişimleri, (b) azalan doğrulama kaybı değişimi gösterilmektedir

%97.66 doğruluk yüzdesine sahip eğitilmiş model, her yeni bulunan normalize edilen karakter bölgesi görüntüsüne uygulanarak, oyma karakterlere sahip dairesel metal cisimlerin sınıflandırma problemi OCR yöntemi ile çözülmüştür. Karakter tanınması sonuçları Şekil 4.12.'de gösterilmektedir.



Şekil 4.12. (a) Normalize edilen karakter kontür bölgesi sonuçları, (b) Karakter tanınması gerçekleştirilmiş görüntü sonuçları

Karakter tanınması işlem süreleri ise, Çizelge 4.3.' te gösterilmektedir. Çizelgede OCR işlemi için uygulanan tüm algoritmalar ve bunların işlem süreleri bulunmaktadır. "n*" ifadesi SWT algoritması sonrası, karakter bölgesi kordinatlara uygulanan matris normalize işlem süresini temsil etmektedir.

Çizelge 4.3. OCR sonrası uygulanan algoritmaların işlem süreleri

ÖRNEK ADI	MSER süresi	SWT + n* süresi	OCR süresi
ANKA	767 ms	358 ms	511 ms
DELPHI	672 ms	167 ms	507 ms
STELLOX	1393 ms	207 ms	477 ms
4U	1288 ms	133 ms	456 ms
A1016	427 ms	178 ms	504 ms
HEARTH	1170 ms	154 ms	525 ms
WONGA	4581 ms	99 ms	461 ms
S11S45	1093 ms	161 ms	468 ms
1118	880 ms	61 ms	504 ms
METZGER	1009 ms	103 ms	513 ms
APLUS	1055 ms	321 ms	461 ms

Dairesel metal cisimlerin sınıflandırma probleminin çözümü için gerçekleştirilen OCR yönteminde kullanılan MSER algoritmasında karakter segmentasyonunun başarısız olduğu örnekleme görüntüleri bulunmaktadır. Bu görüntüler Şekil 4.13.'de gösterilmektedir.



Şekil 4.13. Hatalı MSER sonuçları

Li ve Lu (2012) tarafından gerçekleştirilen çalışmada, MSER algoritmasının doğruluğu için elde ettiği oran %59 olup, karakter bölgelerinin tabela, reklam panosu, uyarı panosu gibi farklı bir arka plan üzerinde durması problemini kolaylaştırmıştır. Dairesel metal cisimlerin üzerindeki karakterler ise oyulmuş olup, farklı bir arka planı bulunmadığından dolayı, ayrıca bazı karakterlerin çok farklı yazı tiplerinde olması ve bazı karakterlerin oldukça bitişik durumda bulunması, problemi zorlaştırmaktadır. Buna karşılık dairese metal cisimler üzerindeki bazı karakterler, bir birleri ile bitişik olmaması,

ve yeterli yazı tipi büyüklüğüne sahip oldukları durumlarda MSER algoritması oldukça başarılı olmaktadır.

Karakter bölgesi tespiti yaklaşımında, Zhou ve ark. (2017) çalışmasında ve Xie ve ark. (2018) çalışmalarında kullanılan derin öğrenme modellerinin elde ettiği doğruluk oranları sırasıyla % 83.27 ve % 88.7' dir. Dolayısıyla derin öğrenme modellerinin karakter bölgesi tespiti için kullanılması MSER algoritmasından çok daha iyi sonuç verdiği görülmesine karşın, gerçekleştirilen çalışmalarda kompleksitenin fazla oluşu ve yüksek çözünürlüğe sahip görüntüler, karakter yazı tiplerinin farklı bir planda bulunması gibi nedenlerden dolayı kullanılmamıştır.

4.8. Sharpen Filtresi Görüntülerinin ESA ile Doğrudan Sınıflandırılması

Bu aşamada, ön karakter bölgesi çıkarılan görsele Sharpen Filtresi (SF) uygulandıktan sonra elde edilen görüntüler, yeniden tasarlanan bir ESA model ile, üzerinde bulunan karakterlere göre sınıflandırma işlemi gerçekleştirilmiştir.

Bu sınıflandırma ile MSER algoritması sonucu başarılı bir şekilde bulunamayan karakterlerin sınıflandırma sonucunu olumsuz etkilemesinden dolayı üçüncü bir sınıflandırma yaklaşımı uygulanmıştır. Bu yaklaşımda, SF sonrası elde edilen görselleri doğrudan ESA modeli için sınıf (label) bilgisi haline getirilmiş ve toplamda her bir dairesel metal cisim için 11 adet sınıf bilgisi elde edilmiştir. Oluşturulan sınıf bilgisi Çizelge 4.4.' te gösterilmektedir. Bu sınıf bilgileri dairesel metal cisimler üzerinde bulunan tam karakter isimleridir.

Çizelge 4.4. Yeni ESA modeli için oluşturulan sınıf bilgileri

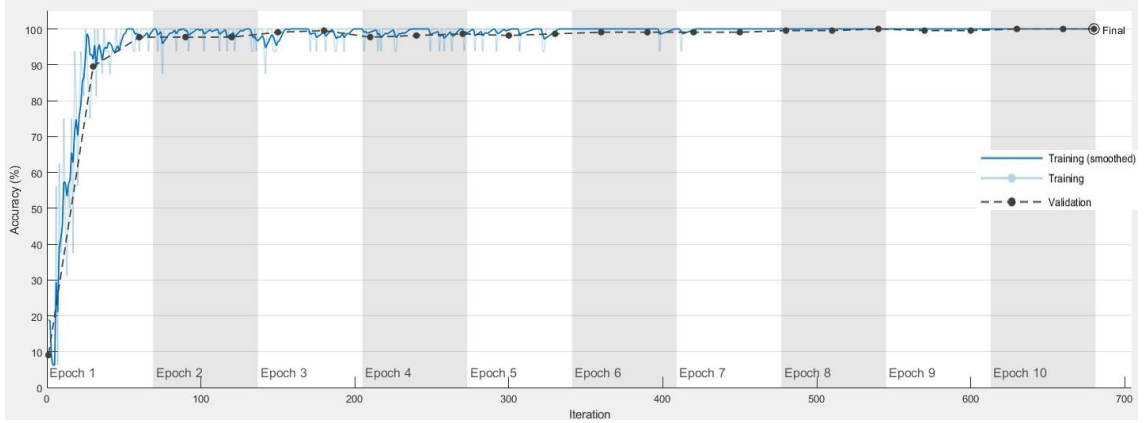
ÖRNEK ADI	Belirlenen Sınıf İsimleri
ANKA	ANKA0916
DELPHI	DELPHIAK42L
STELLOX	STELLOX
4U	4U0916
A1016	A1016
HEARTH	HEARTHBUSS09162975800
WONGA	WONGA0716
S11S45	S11S45
1118	1118
METZGER	METZGER770761018
APLUS	APLUS1218

ESA modelinin eğitim için kullanılan veri seti SF görüntülerin çoğaltılması ile elde edilmiştir. Bu çoğaltma işlemi, ikili tersleme, x ve y eksenlerinde ters çevirme, döndürme, gürültü ekleme, histogram eşitleme, adaptif histogram eşitleme, görüntü yoğunluk değerlerini değiştirme işlemlerinden oluşmaktadır. Veri seti 1100 adet eğitim görüntüsünden, 220 adet ise doğrulama görüntüsünden oluşmaktadır. Yeni oluşturulan ESA modelinin katman detayları Şekil 4.14.'de gösterilmektedir.

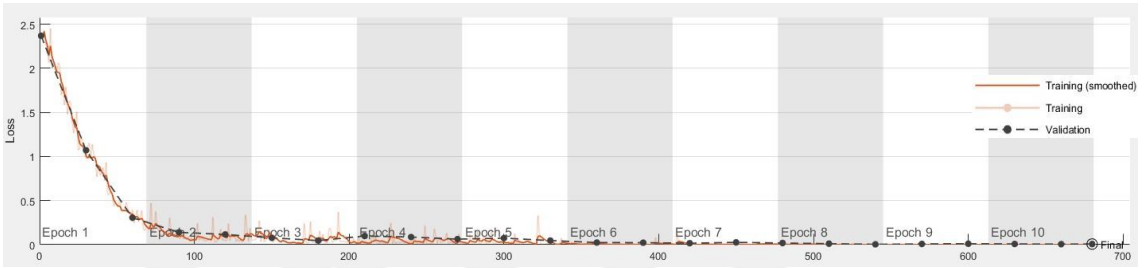
1	imageinput	239x239 images with 'same' normalization	Image Input	239x239x1	-
2	conv_1	32 3x3x1 convolutions with stride [1 1] and padding 'same'	Convolution	239x239x32	Weights 3x3x1x32 Bias 1x1x32
3	batchnorm_1	Batch normalization with 32 channels	Batch Normalization	239x239x32	Offset 1x1x32 Scale 1x1x32
4	relu_1	ReLU	ReLU	239x239x32	-
5	maxpool_1	3x3 max pooling with stride [2 2] and padding [0 0 0 0]	Max Pooling	119x119x32	-
6	conv_2	32 3x3x32 convolutions with stride [1 1] and padding 'same'	Convolution	119x119x32	Weights 3x3x32x32 Bias 1x1x32
7	batchnorm_2	Batch normalization with 32 channels	Batch Normalization	119x119x32	Offset 1x1x32 Scale 1x1x32
8	relu_2	ReLU	ReLU	119x119x32	-
9	maxpool_2	3x3 max pooling with stride [2 2] and padding [0 0 0 0]	Max Pooling	59x59x32	-
10	conv_3	32 3x3x32 convolutions with stride [1 1] and padding 'same'	Convolution	59x59x32	Weights 3x3x32x32 Bias 1x1x32
11	batchnorm_3	Batch normalization with 32 channels	Batch Normalization	59x59x32	Offset 1x1x32 Scale 1x1x32
12	relu_3	ReLU	ReLU	59x59x32	-
13	maxpool_3	3x3 max pooling with stride [2 2] and padding [0 0 0 0]	Max Pooling	29x29x32	-
14	conv_4	32 3x3x32 convolutions with stride [1 1] and padding 'same'	Convolution	29x29x32	Weights 3x3x32x32 Bias 1x1x32
15	batchnorm_4	Batch normalization with 32 channels	Batch Normalization	29x29x32	Offset 1x1x32 Scale 1x1x32
16	relu_4	ReLU	ReLU	29x29x32	-
17	maxpool_4	3x3 max pooling with stride [2 2] and padding [0 0 0 0]	Max Pooling	14x14x32	-
18	conv_5	32 3x3x32 convolutions with stride [1 1] and padding 'same'	Convolution	14x14x32	Weights 3x3x32x32 Bias 1x1x32
19	batchnorm_5	Batch normalization with 32 channels	Batch Normalization	14x14x32	Offset 1x1x32 Scale 1x1x32
20	relu_5	ReLU	ReLU	14x14x32	-
21	maxpool_5	3x3 max pooling with stride [2 2] and padding [0 0 0 0]	Max Pooling	6x6x32	-
22	conv_6	32 3x3x32 convolutions with stride [1 1] and padding 'same'	Convolution	6x6x32	Weights 3x3x32x32 Bias 1x1x32
23	batchnorm_6	Batch normalization with 32 channels	Batch Normalization	6x6x32	Offset 1x1x32 Scale 1x1x32
24	relu_6	ReLU	ReLU	6x6x32	-
25	maxpool_6	3x3 max pooling with stride [2 2] and padding [0 0 0 0]	Max Pooling	2x2x32	-
26	fc	11 fully connected layer	Fully Connected	1x1x11	Weights 11x128 Bias 11x1
27	softmax	softmax	Softmax	1x1x11	-
28	classoutput	classification layer with 1109 and 10 other classes	Classification Output	-	-

Şekil 4.14. SF görüntülerinin doğrudan sınıflandırılması için tasarlanan ESA modelinin katman detayları

Eğitim süreci için ayarlanan giriş görüntüsü 239x239 çözünürlüğünde gri seviye görüntüler olup, tasarlanan modele göre her bir katmanın uygulanması sonucu, nöronlara 2x2x32 çözünürlük bilgisi ulaştırılmıştır. Nöronların sonucunda 11 adet sınıf bilgisi tahmin oranları elde edilerek, en yüksek orana sahip olan ifade tahmin edilen veya karakteri tanınan bilgisini vermektedir. Eğitim sonucunda minimum doğrulama kaybı 0 oranında, doğruluk yüzdesi %100, doğrulanmış doğruluk yüzdesi ise %100 olarak elde edilmiştir. Doğruluk yüzdeleri ve kayıp oranlarını içeren eğitim süreci grafikleri Şekil 4.15.'de gösterilmektedir.



(a)



(b)

Şekil 4.15. Eğitim sürecinde (a) artan doğruluk yüzdesi ve doğrulama doğruluk yüzdesi değişimleri, (b) azalan doğrulama kaybı değişimi gösterilmektedir

Eğitim sonucunda elde edilen %100'lük doğrulama değerli ESA modeli ile SF görüntülerin sınıflandırma işlemi gerçekleştirilmiştir. Modelin tahmin etme süresi ve tahmin edilen sınıf isimleri Çizelge 4.5.'te gösterilmektedir.

Çizelge 4.5. Yeni ESA modelinin tahmin süreleri ve tahmin edilen sınıf isimleri

ÖRNEK ADI	SF süresi	ESA Tahmin süreleri	Tahmin Edilen Sınıf İsimleri
ANKA	24 ms	168 ms	ANKA0916
DELPHI	33 ms	167 ms	DELPHIAK42L
STELLOX	36 ms	168 ms	STELLOX
4U	31 ms	170 ms	4U0916
A1016	24 ms	169 ms	A1016
HEARTH	23 ms	176 ms	HEARTHBUSS09162975800
WONGA	14 ms	167 ms	WONGA0716
S11S45	16 ms	174 ms	S11S45
1118	40 ms	169 ms	1118
METZGER	25 ms	166 ms	METZGER770761018
APLUS	24 ms	170 ms	APLUS1218

Çizelge 4.6.' da dairesel metal cisimlerin sınıflandırılması problemi için önerilen yöntemlerin işlem süreleri ve doğruluk oranları gösterilmektedir. Her iki yöntem içinde yerini belirleme yada daire çaplarının bulunması işlemi ön aşama olarak uygulanmıştır.

Çizelge 4.6. Uygulanan yöntemlerin karşılaştırılması

Uygulanan Yöntemler	Ortalama İşlem Süreleri	Doğruluk ve Eşleşme Oranları
Daire Çaplarının Ölçümü	2.99 s	0.9812
XOR Şablon Eşleştirme Yöntemi	69.72 ms	0.9398
OCR Yöntemi	1.596 s	0.9766
SF Görüntülerinin Sınıflandırılması	195.81 ms	1.0000

Buna göre dairesel metal cisimlerinin ortalama olarak elde edilen, XOR şablon eşleştirme yöntemi için işlem süresi toplamda yaklaşık olarak 3.059 saniye sürmekte, OCR yöntemi için toplamda 4.586 saniye sürmekte, SF görüntülerinin sınıflandırılması yöntemi için ise toplamda 3.185 saniye sürmektedir. Elde edilen işlem sürelerinin azaltılması için gerçekleştirilmesi gerekenler, öneriler bölümünde değinilmiştir.

5. SONUÇLAR VE ÖNERİLER

5.1. SONUÇLAR

Son yıllarda görüntü işleme tabanlı kalite kontrol sistemleri imalat yapan işletmelerin tercih ettiği sistemlerdir. Bu çalışma da görüntü işleme ile oyma karakterlere sahip metal cisimlerin kalite kontrollerini gerçekleştiren bir sistem olarak gerek işletmelere gerekse literatüre katkı sağlayacaktır.

Oyma karakterlere sahip dairesel metal cisimlerinin sınıflandırma problemi için uygulanan XOR şablon eşleştirme yaklaşımı için ön işlem olarak sırasıyla, LoG filtresi, Canny kenar bulma operatörü ve dairesel Hough dönüşümü (DHD) algoritması uygulanmıştır. Daire belirleme doğruluk oranı 0.9812 olarak tespit edilmiş ve en küçük daire çapına sahip görsel ön karakter bölgesi olarak, ham görselden ayrıştırılmıştır. Daha sonra, ön karakter bölgesi görselinin Daughman Rubber Sheet Modeli (DRSM) çıkarımı gerçekleştirilmiştir. Elde edilen DRSM görseline sırasıyla Gabor Filtre (GF) algoritması, Ortalama Mutlak Sapma (OMS) ile özellik çıkarımı ve ikili dönüşüm uygulanarak XOR şablon eşleştirme işlemi için hazır hale getirilmiştir. XOR şablon eşleştirme işlemi sonrasında ise, 69.72 ms ortalama işlem süresi ve 0.9398 ortalama eşleştirme oranı ile sınıflandırma başarısı elde edilmiştir.

Sınıflandırma probleminin bir diğer bakış açısı olan karakter tanınması yaklaşımı için öncelikle dairesel metal cisimlerin yerini belirleme probleminin çözümünde, DHD algoritması sonrası elde edilen ön karakter bölgesi kullanılmıştır. Daha sonra ön karakter bölgesi görseline sırasıyla, Maksimum Kararlı Ekstremal Bölgeler (MSER) algoritması ve Kontür Genişliği Dönüşümü (SWT) algoritması uygulanarak karakter segmentasyonu gerçekleştirilmiştir.

MSER algoritmasının uygulandığı aşamada, karakter segmentasyonunun başarısız olduğu sonuç görüntüleri bulunmaktadır. Bu görüntüler Bölüm 4'de yer alan Şekil 4.13.'de gösterilmektedir. Karakterlerin oyulmuş durumda bulunması, presleme işlemi neticesinde bütün karakterlerin eşit oranda oyulmuş olmaması, karakterlerin farklı fontlarda ve bitişik bir şekilde oyma durumları bulunduğundan dolayı karakter tanıma başarısında düşüş yaşanmıştır.

MSER ve SWT algoritmalarının uygulanması sonucunda elde edilen karakter görselleri, derin öğrenme sınıflarından olan Evrişimli Sinir Ağları (ESA) algoritması ile karakter tanınması gerçekleştirilmiştir. Optik karakter tanıma (OCR) için son aşamada uygulanan ESA modelinin doğruluk başarısı %97.66 olarak elde edilmiş olup, karakter bölgesi tespitinin doğru yapıldığı görsellerde yüksek bir başarı elde etmiştir. OCR

yaklaşımı için sonuç olarak 1.596 s ortalama işlem süresi ve %97.66 doğruluk yüzdesi ile başarımlarını parametreleri elde edilmiştir.

MSER algoritması sonrası elde edilen karakter görselleri kullanarak gerçekleştirilen sınıflandırma işlemi, bazı dairesel metal cisimlerin sınıflandırılmasında verimsiz kalmasından dolayı, SF uygulaması sonrası elde edilen görüntüler doğrudan ESA algoritması ile, cisimler üzerindeki tüm karakter bilgileri kullanılarak sınıflandırma işlemi gerçekleştirilmiştir. Buna göre SF sonrası elde edilen görüntülerin doğrudan ESA ile sınıflandırma yapılması işleminden %100 başarı elde edilerek XOR yöntemine göre daha iyi sonuç elde edilmiştir.

XOR şablon eşleştirme yöntemi hız olarak, OCR yönteminden çok daha hızlı olsa da, karakter büyüklüklerin değişmesi, dairesel metal cisimlerin farklı bir açı ile kamera görselinin oluşması gibi durumlarda doğruluk oranı düşmektedir. OCR yöntemi için ön aşamada uygulanan MSER ve SWT algoritmaları işlem süresi bakımından fazla zaman almakla birlikte, karakterlerin segmentasyonu ve karakter bölgelerinin tespiti için başarılı sonuçlar vermiştir. Dairesel metal cisimlerin sınıflandırma problemi için ürünlerin farklı bir açıda kamera görselinin oluşması, artabilecek veya değişebilecek geniş ürün gamı potansiyeli, yeni ürünlerin XOR şablon eşleştirme için özellik çıkarımı gerçekleştirilmesi ve artan iş-zaman süreci gibi faktörler düşünüldüğünde, OCR çözümünün başarısı önerilmektedir.

5.2. ÖNERİLER

Uygulanan işlemler CPU çekirdeğinde gerçekleştirilmiş olup, Matlab ortamında derlenmiştir. C++ veya Python dilleri ile derlenmesi durumunda, işlem sürelerinin azalacağı tahmin edilmektedir. Ayrıca opencv veya emgucv kütüphaneleri ile GPU çekirdeği üzerinde işlemlerin gerçekleştirilmesi durumunda işlem süreleri ciddi oranda azalacaktır. OCR yönteminde gerçekleştirilen ESA modelinin eğitilmesi için tensorflow veya keras kütüphanelerinin kullanılması durumunda eğitim süresi veya modelin tahmin etme süresi azalacaktır. Ayrıca ESA modelinin eğitilmesi ve tahmin etme sürelerinin daha fazla azaltılması için bir ASIC olan TPU çekirdeği kullanılabilir.

XOR yöntemi ile elde edilen sınıflandırma sonucu, MSER ve SWT kullanılarak elde edilen karakter bilgilerinin sorgulanması ile birleştirilerek sınıflandırma doğrulaması yapılabilir. Böylece sınıflandırma başarısı artırılabilir.

KAYNAKLAR

- Bala, A. ve Tajinder, K., 2016, Local texton XOR patterns: A new feature descriptor for content-based image retrieval, *Engineering Science and Technology, an International Journal*, 19 (1), 101-112.
- Bell, A. J. ve Sejnowski, T. J., 1997, The Independent Components of Natural Scenes are Edge Filters, *Vision Research*, 37 (23), 3327-3338.
- Bishop, C. M., 2006, *Pattern Recognition and Machine Learning (Information Science and Statistics)*, Springer-Verlag, 115-116.
- Canny, J., 1986, A Computational Approach to Edge Detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8 (6), 679-698.
- Chapelle, O., Haffner, P. ve Vapnik, V. N., 1999, Support vector machines for histogram-based image classification, *Ieee Transactions on Neural Networks*, 10 (5), 1055-1064.
- Chawla, S. ve Oberoi, A., 2011, A Robust Algorithm for Iris Segmentation and Normalization using Hough Transform, *Global Journal of Business Management and Information Technology*, 69-76.
- Clausi, D. A. ve Jernigan, T. E., 2000, Designing Gabor filters for optimal texture separability, *Pattern Recognition*, 33 (11), 1835-1849.
- Cohen, G., Afshar, S., Tapson, J. ve van Schaik, A., 2017, EMNIST: an extension of MNIST to handwritten letters, *International Joint Conference On Neural Networks (IJCNN)*, 2921-2926.
- Connell, S. D. ve Jain, A. K., 2001, Template-based online character recognition, *Pattern Recognition*, 34 (1), 1-14.
- D'Albe, E. E. F., 1920, The Optophone: An Instrument for Reading by Ear, *Nature*, 105 (2636), 295-296.
- Dalal, N. ve Triggs, B., 2005, Histograms of Oriented Gradients for Human Detection, *Proceedings - 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005*, 1 (1), 886-893.
- Daugman, J., 1997, Neural image processing strategies applied in real-time pattern recognition, *Real-Time Imaging*, 3 (3), 157-171.
- Daugman, J., 2003, The importance of being random: statistical principles of iris recognition, *Pattern Recognition*, 36 (2), 279-291.
- Donoser, M. ve Bischof, H., 2006, Efficient Maximally Stable Extremal Region (MSER) Tracking, *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1, 553-560.
- Driss, S. B., Soua, M., Kachouri, R. ve Akil, M., 2017, A comparison study between MLP and convolutional neural network models for character recognition, *Real-Time Image and Video Processing 2017*.
- Epshtein, B., Ofek, E. ve Wexler, Y., 2010, Detecting Text in Natural Scenes with Stroke Width Transform, *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2963-2970.
- Erdinc Kocer, H. ve Kursat Cevik, K., 2011, Artificial neural networks based vehicle license plate recognition, *Procedia Computer Science*, 3, 1033-1037.
- Federico, P., Marzal, A., Martin, S. ve Castro, R. R.-G. a. M. J., 2009, A Template-based Recognition System for On-line Handwritten Characters, *Journal of Information Science and Engineering*, 25, 779-791.
- Fukushima, K., 1980, Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position, *Biological Cybernetics*, 36 (4), 193-202.
- Gabor, D., 1946, Theory of communication. Part 3: Frequency compression and expansion, *Journal of the Institution of Electrical Engineers - Part III: Radio and Communication Engineering*, 93 (26), 445-457.
- Gonzalez, A., Bergasa, L. M., Yebes, J. J. ve Bronte, S., 2012, Text Location in Complex Images, *21st International Conference on Pattern Recognition (ICPR 2012)*, 617-620.

- Hanif, S. M. ve Prevost, L., 2009, Text Detection and Localization in Complex Scene Images using Constrained AdaBoost Algorithm. 2009 10th International Conference on Document Analysis and Recognition: 1-5.
- Havlicek, J. P., Havlicek, J. W. ve Bovik, A. C., 1997, The analytic image, *Proceedings of International Conference on Image Processing*, 446-449.
- He, K., Gkioxari, G., Dollár, P. ve Girshick, R., 2017, Mask R-CNN, *Proceedings of the IEEE International Conference on Computer Vision*, 1, 2980-2988.
- Heikkilä, M., Pietikainen, M. ve Schmid, C., 2009, Description of interest regions with local binary patterns, *Pattern Recognition*, 42 (3), 425-436.
- Hinton, G. E., Osindero, S. ve Teh, Y. W., 2006, A Fast Learning Algorithm for Deep Belief Nets, *Neural Computation*, 18 (7), 1527-1554.
- Hough, P. V. C., 1962, General Purpose Visual Input for a Computer, *Brookhaven National Laboratory, Upton, N. Y.*, 99, 323-334.
- Huizhong, C., Tsai, S. S., Schroth, G., Chen, D. M., Grzeszczuk, R. ve Girod, B., 2011, Robust text detection in natural images with edge-enhanced maximally stable extremal regions, *Proceedings - International Conference on Image Processing, ICIP*, 2609-2612.
- Ioffe, S. ve Szegedy, C., 2015, Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift, *CoRR*, 448-456.
- Jailin Reshma, A., Jenushma James, J., Kavya, M. ve Saravanan, M., 2016, An overview of character recognition focused on offline handwriting, *ARPN Journal of Engineering and Applied Sciences*, 11 (15), 9372-9378.
- Keysers, D., Deselaers, T., Gollan, C. ve Ney, H., 2007, Deformation models for image recognition, *IEEE Trans Pattern Anal Mach Intell*, 29 (8), 1422-1435.
- Kingma, D. P. ve Ba, J. L., 2015, Adam: A method for stochastic gradient descent, *ICLR: International Conference on Learning Representations*.
- Krizhevsky, A., Sutskever, I. ve Hinton, G. E., 2012, ImageNet Classification with Deep Convolutional Neural Networks, *ImageNet Classification with Deep Convolutional Neural Networks*, 1097-1105.
- Lam, W. C. Y. ve Yuen, S. Y., 1996, Efficient technique for circle detection using hypothesis filtering and Hough transform, *Iee Proceedings-Vision Image and Signal Processing*, 143 (5), 292-300.
- LeCun, Y., Bottou, L., Bengio, Y. ve Haffner, P., 1998, Gradient-based learning applied to document recognition, *Proceedings of the IEEE*, 86 (11), 2278-2324.
- LeCun, Y., Bengio, Y. ve Hinton, G., 2015, Deep learning, *Nature*, 521 (7553), 436-444.
- Lee, H., Grosse, R., Ranganath, R. ve Ng, A. Y., 2009, Convolutional Deep Belief Networks for Scalable Unsupervised Learning of Hierarchical Representations, *Proceedings of the 26th Annual International Conference on Machine Learning - ICML '09*, 609-616.
- Lee, T. S., 1996, Image representation using 2D gabor wavelets, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18 (10), 959-971.
- Li, Y. ve Lu, H., 2012, Scene Text Detection via Stroke Width, *Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012)*, 681-684.
- Lindeberg, T., 1993, Detecting Salient Blob-Like Image Structures and Their Scales with a Scale-Space Primal Sketch - a Method for Focus-of-Attention, *International Journal of Computer Vision*, 11 (3), 283-318.
- Long, J., Shelhamer, E. ve Darrell, T., 2015, Fully Convolutional Networks for Semantic Segmentation. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 3431-3440.
- Ma, J., Shao, W., Ye, H., Wang, L., Wang, H., Zheng, Y. ve Xue, X., 2018, Arbitrary-Oriented Scene Text Detection via Rotation Proposals, *IEEE Transactions on Multimedia*, 20 (11), 3111-3122.
- Ma, L., Wang, Y. ve Tan, T., 2002, Iris recognition based on multichannel Gabor filtering. *Proceedings of the International Conference on Asian Conference on Computer Vision*, 279-283.
- Marr, D. ve Hildreth, E., 1980, Theory of Edge-Detection, *Proceedings of the Royal Society Series B-Biological Sciences*, 207 (1167), 187-217.

- Matas, J., Chum, O., Urban, M. ve Pajdla, T., 2004, Robust wide-baseline stereo from maximally stable extremal regions, *Image and Vision Computing*, 22 (10), 761-767.
- Mikolajczyk, K., Tuytelaars, T., Schmid, C., Zisserman, A., Matas, J., Schaffalitzky, F., Kadir, T. ve Gool, L. V., 2005, A Comparison of Affine Region Detectors, *International Journal of Computer Vision*, 65 (1-2), 43-72.
- Müller, K. R., Mika, S., Ratsch, G., Tsuda, K. ve Scholkopf, B., 2001, An Introduction to Kernel-Based Learning Algorithms, *Ieee Transactions on Neural Networks*, 12, 181-201.
- Neumann, L. ve Matas, J., 2012, Real-Time Scene Text Localization and Recognition, *2012 IEEE Conference on Computer Vision and Pattern Recognition*, 3538-3545.
- Ojala, T., Pietikainen, M. ve Harwood, D., 1996, A comparative study of texture measures with classification based on feature distributions, *Pattern Recognition*, 29 (11), 51-59.
- Oppenheim, A. V., Lim, J. S. ve Curtis, S. R., 1983, Signal Synthesis and Reconstruction from Partial Fourier-Domain Information, *Journal of the Optical Society of America*, 73 (11), 1413-1420.
- Pratt, W. K., 2001, Digital Image Processing: PIKS Inside, John Wiley and Sons, Inc., New York, NY, USA, 3rd edition, p. 590-595.
- Prewitt, J. M. S., 1970, Object enhancement and extraction, *Picture Processing and Psychopictorics*.
- Qian, S., Liu, H., Liu, C., Wu, S. ve Wong, H. S., 2018, Adaptive activation functions in convolutional neural networks, *Neurocomputing*, 272, 204-212.
- Radzi, S. A. ve Khalil-Hani, M., 2011, Character Recognition of License Plate Number Using Convolutional Neural Network, Berlin, Heidelberg, 45-55.
- Redmon, J. ve Farhadi, A., 2017, YOLO9000: Better, Faster, Stronger, *30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, 6517-6525.
- Redmon, J. ve Farhadi, A., 2019, YOLOv3: An Incremental Improvement, *CoRR*.
- Roberts, L. G., 1963, Machine perception of three-dimensional solids, 315, *Lexington, Mass., M.I.T. Lincoln Laboratory*, p. 40.
- Salembier, P. ve Garrido, L., 2000, Connected operators based on region-tree pruning strategies. Proceedings 15th International Conference on Pattern Recognition. ICPR-2000, 367-370.
- Schantz, H. F., 1982, History of OCR, Optical Character Recognition, Recognition Technologies Users Association, p. 14.
- Schapire, R. E. ve Singer, Y., 1999, Improved boosting algorithms using confidence-rated predictions, *Machine Learning*, 37 (3), 297-336.
- Scott T. Acton, P. Havlicek ve Alan Conrad Bovik, 2001, Oriented Texture Completion by AM-FM Reaction-Diffusion, *IEEE TRANSACTIONS ON IMAGE PROCESSING*, 10 (6), 885-896.
- Shivakumara, P., Phan, T. Q. ve Tan, C. L., 2011, A Laplacian Approach to Multi-Oriented Text Detection in Video, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33 (2), 412-419.
- Smereka, M. ve Duleba, I., 2008, Circular object detection using a modified Hough transform, *International Journal of Applied Mathematics and Computer Science*, 18 (1), 85-91.
- Smith, A. R., 1978, Color Gamut Transform Pairs, *ACM Siggraph Computer Graphics*, 12 (3), 12-19.
- Sobel, I. ve Feldman, G., 1968, An Isotropic 3x3 Image Gradient Operator, *Presentation at Stanford A.I. Project 1968*, 271-272.
- Subrahmanyam, M., Wu, Q. M. J., Maheshwari, R. P. ve Balasubramanian, R., 2013, Modified color motif co-occurrence matrix for image indexing and retrieval, *Computers & Electrical Engineering*, 39 (3), 762-774.
- Tan, T. N., 1995, Texture edge detection by modelling visual cortical channels, *Pattern Recognition*, 28 (9), 1283-1298.
- Tian, Y., Yang, G., Wang, Z., Wang, H., Li, E. ve Liang, Z., 2019, Apple detection during different growth stages in orchards using the improved YOLO-V3 model, *Computers and Electronics in Agriculture*, 157, 417-426.

- Tisse, C., Martin, L., Torres, L. ve Robert, M., 2002, Person identification technique using human iris recognition, *Proceedings of Vision Interface*, 294-299.
- Turner, M. R., 1986, Texture-Discrimination by Gabor Functions, *Biological Cybernetics*, 55 (2-3), 71-82.
- Vintsyuk, T. K., 1968, Speech discrimination by dynamic programming, *Cybernetics and System Analysis*, 4 (1), 52-57.
- Xie., E., Zang., Y., Shao., S., Yu., G., Yao., C. ve Li., G., 2018, Scene Text Detection with Supervised Pyramid Context Network, *CoRR*.
- Ylajaaski, A. ve Kiryati, N., 1994, Adaptive Termination of Voting in the Probabilistic Circular Hough Transform, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16 (9), 911-915.
- Zhang, B., Gao, Y., Zhao, S. ve Liu, J., 2010, Local derivative pattern versus local binary pattern: face recognition with high-order local pattern descriptor, *IEEE Trans Image Process*, 19 (2), 533-544.
- Zhao, X., Lin, K. H., Fu, Y., Hu, Y., Liu, Y. ve Huang, T. S., 2011, Text from corners: a novel approach to detect text and caption in videos, *IEEE Trans Image Process*, 20 (3), 790-799.
- Zhou, X., Yao, C., Wen, H., Wang, Y., Zhou, S., He, W. ve Liang, J., 2017, EAST: An efficient and accurate scene text detector, 2642-2651.

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı Soyadı : Mahmut Sami YASAK
Uyruğu : T.C.
Doğum Yeri ve Tarihi : Selçuklu / 01.12.1992
Telefon : +90 (537) 705 5543
Faks :
e-mail : mhmtsmiyasak@gmail.com

EĞİTİM

Derece	Adı, İlçe, İl	Bitirme Yılı
Lise	: Konevi Anadolu Lisesi, Selçuklu, Konya	2010
Üniversite	: KTO Karatay Üniversitesi, Karatay, Konya	2014
Yüksek Lisans	: Selçuk Üniversitesi, Selçuklu, Konya	
Doktora	:	

İŞ DENEYİMLERİ

Yıl	Kurum	Görevi
2019	Rotechvision	Ar-Ge Mühendisi
2016	Tescom Telekom	Ar-Ge Mühendisi
2015	Stabiley Mühendislik	Ar-Ge Mühendisi

UZMANLIK ALANI

Stabilizasyon Sistemleri / 6-3 6-6 Stewart Platform
 Endüstriyel Otamasyon Sistemleri
 Tıbbi ve Endüstriyel Görüntü İşleme
 Makine Öğrenmesi, Derin Öğrenme Algoritmaları

YABANCI DİLLER

İngilizce

YAYINLAR

Yasak, M.S., Kocer, H. E., Solving the Classification Problem of Circular Metal Objects with Engraved Characters by Image Processing Methods, Konya Journal of Engineering Sciences, 2019 (basıma kabul edilmiştir, link: <https://dergipark.org.tr/konjes/accepted-articles>)