





Haploid ve Diploid Mısırların Aktif Öğrenme Yaklaşımı ile Sınıflandırılması

Active Learning-Based Classification of Haploid and Diploid Maize Seeds

¹Ali GÜNEŞ , ²Emrah DÖNMEZ 

¹Limunis Information Technologies, 17 Green Lanes, N169BS, London / United Kingdom

²Bandırma Onyedi Eylül Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi, 10200, Bandırma/Balıkesir, Türkiye

¹aligunes@limunis.com, ²emrahdonmez@bandirma.edu.tr

Araştırma Makalesi/Research Article

ARTICLE INFO

Article history

Received : 20 January 2023
Accepted : 12 April 2023

Keywords:

Active Learning, Computer Assisted Agriculture, Maize Seed Classification, Support Vector Machine

ABSTRACT

Among agricultural products, the maize plant is shown as the main food source. The breeding process is one of the most important issues for maize plants. In breeding processes, determining the haploid and diploid types of seeds is of great importance. Traditionally, expert analysis is used to detection of haploid and diploid corn types in food commodities. Expert analysis is time-consuming, subjective, and expensive. Computer vision-based solutions are very reasonable requirements agricultural applications. This study presents a novel imaging system that utilizes active learning to classify haploid and diploid corns. Unlike traditional techniques, the system utilizes an active learning method to identify and label informative samples for the classifier. As a result, the proposed system allows for training a support vector classifier with improved accuracy using a smaller sample size.

© 2023 Bandırma Onyedi Eylül University, Faculty of Engineering and Natural Science. Published by Dergi Park. All rights reserved.

MAKALE BİLGİSİ

Makale Tarihleri

Gönderim : 20 Ocak 2023
Kabul : 12 Nisan 2023

Anahtar Kelimeler:

Aktif Öğrenme, Bilgisayar Destekli Tarım, Destek Vektör Makinesi, Mısır Tohumu Sınıflandırma

ÖZET

Tarım ürünleri arasında mısır bitkisi başlıca besin kaynakları içerisinde gösterilmektedir. Mısır bitkilerinde karşılaşılan önemli konulardan birisi ıslah sürecidir. Islah çalışmalarında tohumların haploid ve diploid tiplerinin belirlenmesi büyük önem taşımaktadır. Geleneksel olarak, gıda ürünlerinde haploid ve diploid mısır türlerinin saptanması için uzman analizi kullanılır. Uzman analizi, zaman alıcı, subjektif ve pahalıdır. Bilgisayarlı görü tabanlı çözümler, tarımsal uygulamalar için oldukça makul gereksinimlerdir. Bu çalışmada, haploid ve diploid mısırların aktif öğrenme kullanılarak saptanması için bir görüntüleme sistemi önerilmiştir. Geleneksel yaklaşımların aksine, sınıflandırıcı ile ilgili örneklerin algılanması ve etiketlenmesi için aktif bir öğrenme şeması kullanılır. Önerilen sistem kullanılarak, daha az sayıda örnekle daha yüksek doğruluğa sahip bir destek vektör sınıflandırıcı eğitilmiştir.

© 2023 Bandırma Onyedi Eylül Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi. Dergi Park tarafından yayınlanmaktadır. Tüm Hakları Saklıdır.

1. GİRİŞ

Mısır bitkisi dünya üzerinde kullanılan temel besin kaynakları arasında oldukça önemli bir yere sahiptir. Mısırın nitelik ve nicelik olarak yetiştirilmesinde tohumların saf olarak kullanılması organik ve sağlıklı mısır yetiştiriciliği için oldukça önemli bir husustur. Tohumların saflaştırılması amacıyla öncelikle mısır tohumlarının haploid ve diploid olarak ayrıştırılması gerekmektedir. Günümüzde mısır tohumlarının ayrıştırılma işlemi manuel olarak zirai tarım uzmanları tarafından gerçekleştirilmektedir. Bu işlemin elle yapılıyor olması oldukça zaman alıcı, masraflı, subjektif ve emek yoğun bir süreç gerektirmektedir. Diğer taraftan elle yapılıyor olsa dahi haploid ve diploid mısır tohumu ayrıştırmasında hatalı kararlar verilebilmektedir. Bu sorunların üstesinden gelebilmek amacıyla bilgisayar destekli tarım sistemleri geliştirilmektedir.

Son zamanlarda bilgisayar destekli tarım sistemleri üzerine kayda değer sayıda çalışma gerçekleştirilmiş ve günümüzde de bu kapsamda çalışmalar yapılmaya devam etmektedir. Bilgisayar destekli çözümler, tarım maliyetlerin düşürülmesi, hızlı sonuç alınması, hataların asgari seviyeye indirgenmesi, iş yükünün azaltılması ve güvenliğin artırılması gibi hususlarda oldukça önemli avantajlar sağlamaktadır. Bilgisayar destekli tarım sistemlerinin daha da yaygınlaştırılmasıyla insanoğlunun özellikle tarımda zayıf kalınan bölgeler başta olmak üzere tarımsal ürünlere olan ihtiyaçların yeterli seviyede karşılanması amaçlanmaktadır. Diğer bir husus olarak bu sistemlerin yaygınlaşması ile organik tarım ürünü üretiminde de hızlanmanın sağlanacağı ve daha sağlıklı ürünlerin daha düşük maliyetlerle üretileceği öngörülmektedir. Bilgisayar destekli tarım sistemlerinin haploid ve diploid mısır tohumlarının ayrıştırılmasında da kullanılabilmesi ve bu yolla tohum saflaştırma süreçlerine katkı sağlanacağı düşünülmektedir. Bu amaçla, özellikle son zamanlarda mısır tohumu ayrıştırılması üzerine önemli katkılar sağlayan çalışmalar yapılmıştır.

Gerçekleştirilen çalışmalarda haploid ve diploid mısır tohumu sınıflandırılması çalışmalarında başlarda DNA analizi, tohum yağı analizi, bitkinin yaprak, kök vb. bileşenlerinin biyometrik olarak analizi gibi bitki içerik veya özelliklerine dayalı çalışmalar gerçekleştirilmiştir. Daha sonraları ise mısır tohum görüntülerinin yakın kızılötesi, çoklu/hiper spektral ve floresan görüntüleme gibi özel görüntüleme donanımlarıyla elde edilmesi ve tohumların bu görüntülere göre sınıflandırılmasına yönelik çalışmalar yürütülmüştür. Mevcut güncel çalışmalarda ise artık makine öğrenmesi yöntemleri ile mısır tohumlarına ait özelliklerin analiz edilmesiyle haploid ve diploid ayrımı yapılmaya başlanmıştır. Görüntü tabanlı olarak gerçekleştirilen çalışmaların çoğunda mısır tohumu görüntüsü üzerindeki embriyo bölgesinde yer alan ve R1-nj renklenmesi olarak isimlendirilen ayırt edici özellik kullanılarak haploid ve diploid mısır tohumlarının sınıflandırılması işlemi gerçekleştirilmiştir. Son zamanlarda yapılan çalışmalarda göze çarpan kayda değer bir diğer durum ise özel görüntüleme donanımları kullanarak mısır tohumlarına ait özneliklerin çıkarılması yerine sadece basit bir kamera ile de mısır görüntülerine ait özneliklerin elde edilmesidir. Böylece özel donanımlara olan ihtiyaç azaltılarak daha ucuz maliyetli mısır tohumu sınıflandırma sistemlerinin geliştirilmesine olanak sağlanmaktadır.

Elde edilen özneliklerle sınıflandırma işlemi bilgisayar destekli sistemlerle gerçekleştirmek için makine öğrenmesi algoritmaları geliştirilmiştir. Makine öğrenmesi yöntemleri dijital verilerin sınıflandırılması ve regresyon analizi gibi işlemler için kullanılmaktadır. Makine öğrenmesi yöntemleri üzerinde çalışılacak verilerin yapısına göre seçilir. Verilerin elde edilmesinden sonra tüm veriler üzerindeki ön-tanımlı veya ön işlem ile belirlenmiş özneliklere ait değerler hesaplanır. Bu değerler kararlı bir aralığa oturma kadar özneliklerin güncellenmesine devam edilir. Elde edilen kararlı öznelik değerlerine göre sisteme gelen yeni verinin hangi sınıfa ait olduğu belirlenir. Gelen verilerin doğru bir şekilde sınıflandırılması ise sınıflandırıcı yönteminin başarımını belirler. Bu kapsamda tarım ürünlerine ait veriler de makine öğrenmesi yöntemleriyle sınıflandırılabilir. Mısır tohumlarına ait görüntüler üzerinde elde edilen öznelik değerlerine bağlı olarak haploid ve diploid mısır tohumu ayrıştırılması işlemi gerçekleştirilebilir.

Bu çalışmada, mısır tohumlarının renk moment öznelikleri kullanılarak aktif öğrenme tabanlı yüksek doğruluklu bir sınıflandırıcı önerilmiştir. Sınıflandırıcı olarak Destek Vektör Makinesi kullanılmıştır. Geleneksel sınıflandırma yaklaşımlarından farklı olarak önerilen aktif öğrenme yönteminde minimum sayıda ve en ayırt edici bilgileri taşıyan örnekler ile sınıflandırıcı eğitilmiştir. Böylece en yüksek sınıflandırma başarımının, en az sayıda örneğin etiketlenerek elde edilebilmesine olanak tanınır. Bu makalenin ikinci bölümünde ilişkili çalışmalar sunulmuştur. Çalışmada kullanılan veri-kümesi, öznelik çıkarma, öznelik seçme, sınıflandırıcı ve aktif öğrenme stratejisi üçüncü bölümde sunulmuştur. Deneysel sonuçlar ve elde edilen gözlemler dördüncü bölümde verilmiştir. Beşinci bölümde ise tartışma ve sonuca değinilmiş ve gelecek çalışmalardan söz edilmiştir.

2. İLİŞKİLİ ÇALIŞMALAR

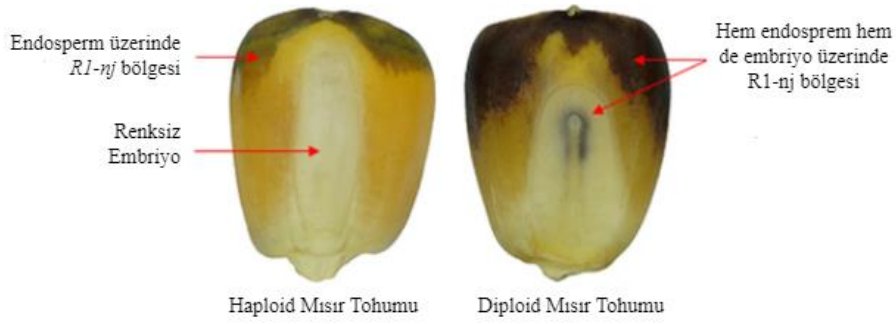
Boote vd. [1], floresan mikro spektroskopisi ve görüntüleme yöntemlerinin bir kombinasyonu ile r1-nj renk belirteci kullanarak mısır tohumlarındaki kendilenmiş haploid farklılığını tespit eden yeni bir floresan tabanlı yöntem önerdiler. Yedi farklı kendilenmiş haploid hatları üzerindeki denemelerde %80 ve beş farklı kendilenmiş haploid hatları üzerindeki denemelerde %90 başarıma ulaştılar. Fuente vd. [2], katlanmış haploid'lerin mısır ıslah programlarındaki önemine değindiler. Çalışmada, R1-nj renk belirteci ile haploid mısır tohumlarının manuel tanımlanmasının emek-yoğun ve zaman alıcı bir süreç olduğu ifade edildi. Önerdikleri yaklaşımda altı farklı kendilenmiş hat ana haploid indükleyici ile çaprazlandı. Her bir hat için tohum örneği manuel olarak sıralandı. VideometerLab 3 adında harici bir yazılım kullanarak spektral görüntüleme yöntemini haploid ve melez tohumları

ayrıt etmek için kullandılar. Çalışmada bu yaklaşımla %50 üzerinde bir doğruluğun sağlanmasının mümkün olduğu ifade edildi. Lin vd. [3], mısır ıslahında haploid tohumların tanımlanmasının oldukça önemli olduğunu söylediler. Çalışmada, yakın-Kızılötesi spektroskopide bulunan yaygın aktarım teknolojisi ile NIR spectrum öznelikleri kullanılarak melez haploidlerin tespitinde tanıma oranını düşüren gereksiz öznelıklar de olduğu ifade edildi. Zero-phase bileşen analizi, partial least squares regression ve back propagation neural network yöntemleri kullanarak tohum spektrum özelliklerinin düşük boyutlu ve tek tip uzayını tasarlamak için toplanan spektrumun iyileştirilmesini sağladılar. Altuntaş vd. [4], in vivo maternal haploid tekniğinde yapılan haploid seçiminin çok önemli olduğunu ancak bu seçimin elle yapılıyor olmasının zaman alıcı olduğu ve başarımı düşürdüğünü ifade etmektedirler. Yaptıkları çalışmada haploid ve diploid görüntülerini orijinal görüntüden bölütlediler. Bölütlenen imge üzerinde baş farklı öznelik çıkarıldı. Bu öznelikleri de her bir veriye ait öznelik vektörünün oluşturularak SVM yöntemi ile sınıflandırılmasında kullandılar. Çalışmalarında haploid tanıma oranının %94.25 ve diploid tanıma oranının %77.91 olduğunu ifade ettiler. Wang vd. [5], hızlı ve doğru haploid çekirdeği tanıma yönteminin haploid ıslah verimliliğini hızlandırmada büyük öneme sahip olduğunu söylediler. Mevcut otomatikleştirilmiş haploid tanıma yaklaşımları sınırlılıklarının üstesinden gelmek ve daha doğru haploid görüntüleme sağlamak için yakın-spektral görüntüleme teknolojisi kullanarak haploid maize çekirdeğini tanılamada daha hızlı ve doğru bir yöntemin varlığını araştırdılar. Çalışma sonucunda haploid maize tanımda hiperspektral görüntüleme teknolojisinin elverişli olduğu söylendi. Yu vd. [6], yakın-kızılötesi spektroskopi teknolojisinin haploid tohumları melez tohumlardan ayırmak için yıkıcı olmayan, hızlı ve düşük maliyetli olmak gibi avantajları olduğunu söylediler. Diğer taraftan, ışık, sıcaklık, nem, yakın-kızılötesi yoğunluğu, tohum aktivitesinin enstrüman ve dinamik değişimi gibi parametreler yüzünden maize tohumlarını yakın-kızılötesi spektrası yüksek boyutlu non-lineer karakteristikler gösterdiği ifade edilmektedir. Çalışmada, haploid mısır tohumu tanıma görevi için doğrusal olmayan bir özellik analizi yöntemi önerildi. Altuntaş vd. [7], katlanmış haploid kullanımının mısır ıslahında verimliliğin artırılmasını sağladığını ifade ettiler. Haploid mısır tohumlarının belirlenmesinde genellikle zirai uzmanlar elle tayin yapmaktadırlar. Bu süreç ise emek-yoğun ve zaman alıcı bir görevdir. Toplamda 3000 adet veri üzerinde yürüttükleri çalışmada transfer öğrenme ile CNN modellerine ait eğitim parametrelerini güncellemiş ve VGG-19 CNN modeli üzerinde %94.22 oranında başarıya ulaşmışlardır. Dönmez [8], AlexNet CNN modeline toplamda 3000 adet etiketli mısır görüntüsünü girdi olarak verilerek modelin tam bağlı 'FC6', 'FC7' ve 'FC8' katmanlarından öznelik çıkarımı yapılmıştır. Çalışmada CNN modelinin son katmanındaki SoftMax sınıflandırıcısı yerine SVM sınıflandırıcısı kullanılmış ve %89.5 doğruluk oranı elde edilmiştir. Dönmez [9], beş farklı CNN modelini öznelik çıkarıcı olarak kullandı. Elde edilen öznelik vektörleri SVM, KNN ve DT (Decision Tree) sınıflandırıcılarıyla sınıflandırılmıştır. Çalışmada ResNet-50 CNN modeli ile SVM sınıflandırıcı kombinasyonunun %91,4 düzeyinde başarımlı sağladığı rapor edilmiştir. Dönmez [10], altı farklı güncel ön-eğitilmiş CNN mimarisini öznelik çıkarıcı olarak kullandı. Çıkarılan öznelik vektörleri arasından MRMR öznelik seçim algoritması kullanılarak 100 adet öznelik seçimi yapıldı. Seçilen öznelikler daha sonra farklı kombinasyonlarda kaynaştırılarak DT, KNN ve SVM sınıflandırıcılarını eğitmek ve test etmek için kullanıldı. Çalışmada literatürdeki pahalı ekipman ve çıktılar analizi zaman alan çalışmaları da karşılaştırarak, geliştirilen yöntemin %96,74 oranında başarımlı ulaştığı ifade edildi. Güneş vd. [11], hiperspektral görüntüleme ile elde edilmiş aflatoksinli incir görüntülerini aktif öğrenme yaklaşımı ile yüksek başarımlı sınıflandırmışlardır. Bu çalışmada minimum sayıda etiketli örnekle, performanslı bir sınıflandırıcı elde edilmiştir. Kalkan vd. [12], bir başka çalışmada aflatoksin ve küf bulaşık incir örneklerini hiperspektral görüntüleme ve geleneksel sınıflandırıcılar ile tespit etmişlerdir. Literatürdeki yöntemlerin çoğu incelendiğinde çoğunlukla maliyeti yüksek tohum analizi sistemlerinin geliştirildiği görülecektir. Kullanılan DNA, yağ analizi vb. moleküler seviyedeki yöntemler veya IR, multi/hiper spektral ve floresan görüntüleme vb. görüntüleme yöntemleri maliyeti yüksek yöntemlerdir. Bu çalışmada literatürde var olan çalışmalardan da ilham alarak, haploid ve diploid mısır tohumlarının sınıflandırma analizi için aktif öğrenme modeli önerilmiştir. Önerilen model maliyet odaklı ve yüksek performanslı bir yöntem olarak literatürdeki çalışmalardan ayrılmaktadır. Geliştirilen yöntemle mısır tohumlarının sınıflandırılması için gelecek vaat eden sonuçlar ortaya koyulmuştur.

3. MATERYAL VE METOTLAR

3.1. Veri Kümesi

Bu çalışmada önerilen yöntemler, uygun laboratuvar koşullarında elde edilen mısır tanelerine ait görüntüler üzerinde test edilmiştir. Veri kümesine ait mısır taneleri Sakarya bölgesinden toplanmış olup, Sakarya Mısır Araştırma Enstitüsündeki konu uzmanlarınca R1-nj renk özneliklerine göre manuel olarak etiketlenmiştir (Şekil 1). Etiketlenen bu veriler daha sonra konvansiyel görüntüleme cihazları kullanılarak dijital ortama aktarılmış ve bu imgeler ön işlemlerden geçirilmiştir [7]. Toplamda 1230 haploid ve 1770 diploid mısır tohumu görüntüsü bulunan veri kümesinden rastgele seçilen 1200 haploid ve 1200 diploid veri kullanılmıştır. Burada amaç farklı sınıflara ait veri sayıları arasında dengeli bir dağılım sağlamaktır.



Şekil 1. R1-nj renk işaretleyicisine göre haploid ve diploid mısır tohumlarının görsel sınıflandırılması.

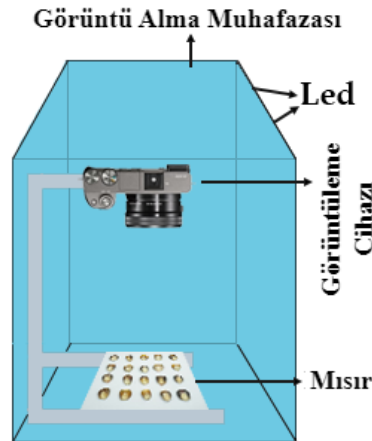
Şekil 2’de veri kümesinden rastgele biçimde seçilmiş diploid (üstte) ve haploid (altta) mısır tohumlarına ait imgeler verilmiştir.



Şekil 2. Örnek mısır tohumu görüntüleri, diploid (üst), haploid (alt)

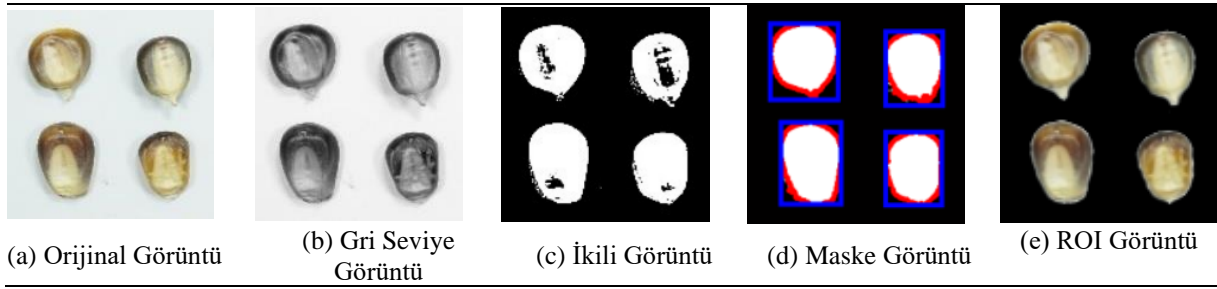
3.2. Görüntüleme Sistemi ve Önışlemler

Bu çalışmada önerilen yöntemler, uygun laboratuvar koşullarında elde edilen mısır tanelerine ait görüntüler üzerinde test edilmiştir. Veri kümesine ait mısır taneleri Sakarya bölgesinden toplanmış olup, Sakarya Mısır Araştırma Enstitüsündeki konu uzmanlarınca R1-nj renk özneliklerine göre manuel olarak etiketlenmiştir. Toplamda 2400 adet mısır tohumu görüntüsünden oluşan veri kümesinde 1200 adet haploid ve 1200 adet diploid tipte mısır tohumu görüntüsü bulunmaktadır. Mısır tanelerine ait görüntülerin alınması, laboratuvarında kurulan görüntüleme sistemi (Şekil 3) yardımıyla yapılmıştır.



Şekil 3. Mısır tanelerinin görüntülenmesi için kurulan görüntüleme sistemi.

Kurulan görüntüleme sisteminde kamera olarak SONY ILCE-6000 dijital görüntüleme cihazı kullanılmıştır. Işık kaynağı olarak görüntüleme sisteminin tavanına yerleştirilen LED ışık kaynakları kullanılmıştır. Mısır taneleri, görüntü düzleminde birbirlerine temas etmeyecek şekilde ve her görüntüde 20 mısır tanesi olacak şekilde görüntülenmiştir. Çekim sürecinde fotoğraf makinası M modunda, 1/125 sn enstantene, 200 ISO, diyafram açıklığı f/7,1, odak uzaklığı 24mm, kamera uzaklığı 18 cm ve 4240 x 2832 çözünürlük ayarlarına kurulmuştur [4]. Toplu halde yirmi adetli kümelerden oluşan mısır tanelerinin bulunduğu görüntüler bölütlenerek tekil mısır tane görüntüleri elde edilmiştir. Elde edilen her bir mısır tanesi görüntüsü RGB renk uzayındadır. Elde edilen görüntüler üzerindeki mısır bölgesini (ROI) tespit etmek ve görüntüler üzerindeki gürültüleri temizlemek amacıyla maskeleme işlemi uygulanmıştır. Bu amaçla her bir mısır görüntüsü eşiklenerek ilgili maske görüntüleri oluşturulmuştur (Şekil 4).



Şekil 4. Haploid ve diploid mısır tanelerine ait maske ve ROI görüntüleri.

Elde edilen maske görüntüleri mısır görüntüleri ile çarpılarak her bir mısır tanesine ait ROI görüntü elde edilmiştir (Şekil 4-e).

3.3. Öznitelik Çıkarma ve Seçme

Günümüzde yapılan sınıflandırma temelli çalışmalarda sınıflandırma başarımını arttıran faktörlerden birisi de ayırtedici bilgiler taşıyan özniteliklerin mümkün olduğunca fazla olmasıdır. Ancak buna bağlı olarak öznitelik sayısı ile örnek sayısı arasındaki ilinti ve sınıflandırıcı karmaşıklığı sınıflandırıcı performansını etkileyen temel unsurlardır. Sınıflandırma karmaşıklığının fazla olması hafıza gereksinimine olan ihtiyacı artırmakla birlikte sınıflandırıcı hızını da düşürmektedir. İdeal bir sınıflandırıcı sistemi için temel gereklilik, ayırt edici bilgiler taşıyan yeterli sayıda öznitelik kümesinin belirlenebilmesidir. Sınıflandırma problemlerinde boyut indirgeme stratejileri temel olarak öznitelik çıkarma ve seçme başlıkları altında toplanmıştır.

Boyut indirgeme yaklaşımlarından birisi olan özniteliklerin belirlenmesi X boyutlu bir örnekler kümesinde yer alan bir öznitelik grubunun doğrusal olan veya olmayan algoritmalar yardımıyla yeni bir Y boyutlu öznitelik grubuna dönüştürülmesidir. Çalışılan probleme bağlı olarak seçilen öznitelik çıkarma yaklaşımları değişebilmektedir. Elde edilen yeni Y boyutlu öznitelik kümesinde halen gereksiz bilgiler taşıyan öznitelikler yer alabilir. Bu nedenle Y boyutlu öznitelik kümesi içerisinde veriye ve probleme bağlı olarak çeşitli öznitelik seçme yöntemleri kullanılarak daha küçük boyutlu ($Y > Z$) yeni bir Z öznitelik kümesi elde edilir. Boyut indirgeme problemlerinde önemli sorulardan birisi de en uygun öznitelik sayısının nasıl belirleneceğidir. Bu konuda herhangi bir kural olmamakla birlikte sistemi geliştiren kişi tarafından problemin yapısına uygun olarak deneysel olarak belirlenir.

Özellik seçme algoritmalarının arkasındaki motivasyon, sorunla en alakalı özelliklerin bir alt kümesini otomatik olarak seçmektir. Özellik seçiminin amacı iki yönlüdür: Alakasız özellikleri veya gürültüyü kaldırarak hesaplama verimliliğini artırmak ve modelin genelleme hatasını azaltmaktır. En temel öznitelik seçme algoritması, sınıflandırma hata oranını en aza indiren bularak özniteliklerin olası her bir alt-öznitelik kümesini test etmektir. Bu şekilde olası tüm öznitelik kümeleri üretilerek değerlendirme ölçütü sağlanana kadar test edilir. Değerlendirme ölçütünün seçimi, algoritmayı büyük ölçüde etkiler ve özellik seçme algoritmalarının üç ana kategorisini (sarmalayıcılar, filtreler ve gömülü sistemler) birbirinden ayıran da bu değerlendirme ölçüleridir.

Ardışık öznitelik seçim algoritmaları, ilk d -boyutlu öznitelik uzayını, $k < d$ olan bir k -boyutlu öznitelik alt uzayına indirgemek için kullanılan bir sıralı arama algoritmaları ailesidir. Düşük karmaşıklığa sahip, başarılı sonuçlar veren ve kolay uygulanabilen yöntemlerdir. Sıralı arama yöntemleri (Sequential Forward Selection, Sequential Backward Selection) sınıflandırıcı performansına bağlı olarak, k boyutta bir öznitelik alt kümesine ulaşılan kadar her seferinde bir özniteliği kaldırır veya ekler. Kayan (floating) varyantlar SFFS ve SBFS ise daha basit SFS ve SBS algoritmalarının uzantıları olarak kabul edilebilir. Kayan algoritmalar, dahil edildikleri (veya hariç tutuldukları) zaman öznitelikleri kaldırmak için ek bir silme veya ekleme adımına sahiptir, böylece daha fazla sayıda öznitelik alt-kümesi kombinasyonu örneklenebilir. Bu adımın koşullu olduğunu ve yalnızca ortaya çıkan öznitelik alt-kümesinin, belirli bir özneliğin çıkarılmasından (veya eklenmesinden) sonra kriter işlevi tarafından "daha iyi" olarak değerlendirilmesi durumunda gerçekleşir [13]. SFFS yöntemine ait algoritma adımları özetlenirse;

SFFS öznitelik seçim yöntemine ait algoritma adımları aşağıdaki gibi özetlenebilir;

Algoritma 1. Sıralı İleri Kayan Arama Algoritma Adımları

1. $Y_0 = \{\emptyset\}$ öznitelik içermeyen küme ile başla
2. En ayırt edici özniteliği ekle

$$x^+ = \operatorname{argmax}_{x \in Y_k} [J(Y_k + x)] \quad Y_k = Y_k + x^+; k = k + 1$$
3. En zayıf özniteliği çıkar $x^- = \operatorname{argmax}_{x \in Y_k} [J(Y_k - x)]$
4. Eğer $J(Y_k - x^-) > J(Y_k)$ $Y_{k+1} = Y_k - x^-; k = k + 1$

Üçüncü adıma git

Değilse

İkinci adıma git

Bir mısır tanesinin haploid veya diploid olduğunu belirleyen ROI embriyo ve endosperm bölgesi mısır tanelerinde boyut, şekil veya dokusal açıdan bir değişikliğe sebep olmamaktadır. Ancak mısır tanelerinin embriyo ve endosperm bölgelerinde mısırın haploid veya diploid tipte olmasına bağlı olarak renk değişimi gözlemlenebilmektedir. Bu nedenle, haploid bir mısır tanesini diploid bir mısır tanesinden ayıran en güçlü öznelik embriyo ve endosperm bölgesindeki R1-nj renklenmesidir. Bu çalışmada, mısır tanelerine ait ayırt edici öznelik olarak embriyo ve endosperm bölgelerindeki renk momentleri kullanılmıştır. Haploid ve diploid mısır sınıflarını ayırt edebilmek için ayrı ayrı Kırmızı (Red), Yeşil(Green), Mavi (Blue) renk kanallarının ortalama renk yoğunlukları, her bir renk kanalına ait standart sapma değerleri, çarpıklık ve kurtosis değerleri öznelik olarak kullanılmıştır. Bu yolla, haploid ve diploid mısır görüntülerinin ayrıştırılması için bir mısır tanesine ait toplam 12 öznelik elde edilmiştir.

Ortalama (5), standart sapma (6), çarpıklık (7) ve kurtosis (8) değerleri ilk dört dereceden renk momentleri olarak renkli bir mısır görüntüsünün karakteristiğini temsil eden ayırt edici özneliklerdir.

$$\mu = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n piksel_{ij} \quad (5)$$

$$\sigma = \sqrt{\frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n (piksel_{ij} - \mu)^2} \quad (6)$$

$$s = \sqrt[3]{\frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n (piksel_{ij} - \mu)^3} \quad (7)$$

$$k = \sqrt[k]{\frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n (piksel_{ij} - \mu)^4} \quad (8)$$

Denklem 5, 6, 7 ve 8 'de σ standart sapmayı, μ ortalamayı, k kurtosisi ve s ise çarpıklığı ifade etmektedir. $m \times n$ çözünürlükteki bir görüntüde $piksel_{ij}$ i . satır ve j . sütundaki ilgili renk kanalında karşılık gelen piksel değerini ifade etmektedir.

Veri kümesinde yer alan tüm haploid ve diploid mısır örnekleri göz önüne alındığında 2400 x 12 ölçekli bir matris elde edilmiştir. Elde edilen 12 adet öznelik, en ayırt edici bilgiler taşıyandan daha az bilgi taşıyana doğru Suquential Floating Forward Selection (SFFS) yöntemiyle sıralanmıştır.

3.4. Aktif Öğrenme (AÖ)

Aktif öğrenme, sınıflandırıcı eğitimi için gerekli olan eğitim örneklerinden en ayırt edici öznelikleri elde edebilmek amacıyla geliştirilen özelleştirilmiş bir öğrenme sürecidir. Aktif öğrenme sisteminin temel amacı, çok sayıda sınıflandırıcıya karşı seçilen girdi sorgularını gerçekleştirmektir. Rastgele örnekleme ile karşılaştırıldığında, aktif öğrenmeyi (AÖ) kullanan örnek seçimi daha ayırt edicidir. Genel bir aktif öğrenme modeli beş birim kullanılarak modellenir (G, Q, S, T, U). G, etiketli eğitim seti T üzerinde eğitilen denetimli bir sınıflandırıcıdır. Q, etiketlenmemiş örneklerden oluşan bir U havuzundan en bilgilendirici etiketlenmemiş örnekleri seçmek için kullanılan bir sorgulama işlevidir. S, herhangi bir etiketlenmemiş U örneğine gerçek sınıf etiketi atayabilen bir gözetmendir. AÖ işlemi, S süpervizörünün her yinelemede Q sorgu fonksiyonu tarafından seçilen en bilgilendirici örnekleri yinelemeli olarak etiketleyerek sistemle etkileşime girdiği yinelemeli bir süreçtir. İlk aşamada, sınıflandırıcı G'nin ilk eğitimi için birkaç etiketli numuneden oluşan bir başlangıç eğitim seti T gereklidir.

AL yaklaşımı ile yapılan sınıflandırma işlemlerinde yüksek doğruluklu sınıflandırma oranları korunurken hatalı sınıflandırma oranları azaltılır. Sınıflandırma problemlerinde yaygın olarak kullanılan temel bir Aktif Öğrenme algoritmasına ait algoritma adımları aşağıdaki gibidir;

Algoritma 2. Temel Aktif Öğrenme Algoritması

1. G sınıflandırıcısı T eğitim kümesi kullanarak eğit.
2. Etiketsiz U örnekler havuzunu G sınıflandırıcısı kullanarak sınıflandır.

Döngü:

3. Sorgu fonksiyonu Q aracılığı ile etiketsiz örnekler U havuzundan seçilir.
4. U havuzundan seçilen örneklere S uzmanı ile etiket atanır.
5. Dördüncü aşamada eğitim kümesi T 'ye etiketlenen örnekler eklenir.
6. G sınıflandırıcısı yeniden eğitilir.

Sonlandır: Ön-tanımlı ölçütler sağlandığında algoritmada çalışması durdurulur.

S uzman yorumu mısır tanelerinin sınıflandırılmasında başlangıç eğitim kümesi T için kullanılacak etiketli örnekler ve sorgu fonksiyonu Q ile seçilen örneklerin etiketlenmesi için kullanılmaktadır. AÖ algoritmaları, literatürde iki kategoriye ayrılır: örneklerin belirsizliği (uncertainty) [14] ve komite temelli sorulama (query by committee) yaklaşımları [15]. Belirsizlik temelli AÖ yaklaşımları, sınıflandırıcının en çok zorlandığı, belirsizliği en yüksek olan etiketsiz örnekleri seçmeye odaklanır. Bu örnekler, sınıflandırıcının belirsiz olduğu ve daha fazla bilgiye ihtiyacı olduğu örneklerdir. Bu nedenle, belirsizlik temelli yaklaşımlar, örneklerin belirsizliği ölçütünü kullanarak etiketleme işlemi için en iyi örnekleri seçmeye çalışır. Belirsizlik ölçütü, bir sınıflandırıcının ne kadar güvenilir olduğunu ölçmek için kullanılan bir terimdir. Bu ölçüt, sınıflandırıcının doğruluk veya hata oranı gibi değişkenlere dayanarak hesaplanabilir. Sınıflandırıcıya bağlı olarak, belirsizlik ölçütü farklı yöntemlerle ifade edilebilir. Örneğin, bir lojistik regresyon sınıflandırıcısı için belirsizlik ölçütü olasılık değerleri arasındaki fark olarak hesaplanabilirken, bir destek vektör makinesi için belirsizlik ölçütü uzaklık değerleri olarak ifade edilebilir. Örneğin, Bayes sınıflandırıcı kullanıldığında belirsizlik ölçütü olarak sınıflar temelinde hesaplanan sonrasal olasılıklar kullanılabilir. Ayrıca, diğer sınıflandırıcılar için entropi, çoklu sınıflandırma, güven aralığı gibi farklı belirsizlik ölçütleri kullanılabilir. Belirsizlik ölçütünün seçimi, sınıflandırıcının özelliklerine ve uygulamanın amacına göre değişebilir. Lewis ve Gale [14], yaptıkları çalışmada, Bayes sınıflandırıcı kullanarak hesaplanan sonrasal olasılıklarını etiketsiz örneklerin belirsizlik ölçütü olarak kullanmışlardır. Bu yöntem, etiketsiz örneklerin sınıflandırılması için en uygun olanların seçilmesini sağlar. Örneğin, iki sınıflı bir problemde henüz sınıflandırma yapılmamış bir etiketsiz örneğin belirsiz kategorisine girebilmesi için örneğin sonrasal olasılık değerlerinin sınıflar temelinde 0.5'e oldukça yakın olması gerekir. Lewis ve Gale, bu ölçütü kullanarak etiketsiz örneklerin belirsizliğini ölçerek, en iyi örnekleri seçmeye çalışmışlardır.

İki sınıflı (binary) bir problem ele alındığında, bir etiketsiz örneğin belirsizlik durumuna takılması için sonrasal olasılık değerlerinin sınıflar temelinde yaklaşık bir hesapla 0.5 olması gerekmektedir. Bu, sınıflandırıcının verilen örnek için her iki sınıf için de benzer olasılık değerleri verdiği anlamına gelir. Bu durumda, sınıflandırıcının verilen örnek için bir sınıfın daha uygun olduğunu belirleyememesi, yani belirsiz olduğu anlamına gelir.

Literatürde belirsizlik ölçütleri arasında en yaygın ve popüler olanı, Shannon [16] tarafından önerilen entropi temelli belirsizlik yaklaşımıdır. Bu yaklaşım, sınıflandırıcının belirsizliğini ölçmek için entropi kavramını kullanır. Entropi, sınıflandırıcının sınıflar arasındaki kararsızlığını ölçen bir ölçüttür. Örneğin, sınıflandırıcının belirsizliği yüksekse, entropi de yüksek olacaktır. Bu nedenle, entropi temelli belirsizlik ölçütü, sınıflandırıcının belirsizliğini ölçerek, en iyi etiketleme için örnekleri seçmeye yarar.

$$\text{Belirsizlik}(x) = \sum_i P_{\theta}(y_i|x) \log P_{\theta}(y_i|x) \quad (9)$$

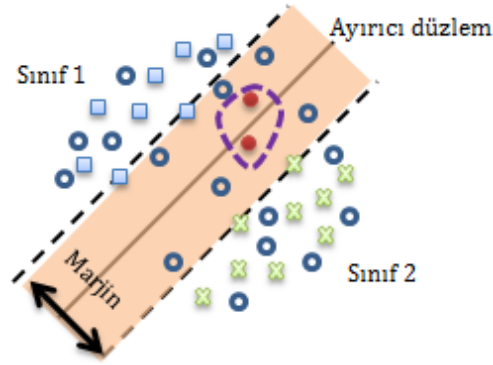
Denklem 9 'da yer alan y_i parametresi, tüm olası etiketleri ifade etmektedir. Bu parametre, verilen örneğin sınıflandırılması için kullanılan etiketi temsil eder. Çok sınıflı problemlerde, y parametresi birden fazla değer alabilir, çünkü örnekler birden fazla sınıfa ait olabilir. Örneğin, y parametresi $[0,1,0,0,1]$ şeklinde olabilir, bu da örneğin ikinci ve beşinci sınıfa ait olduğunu gösterir. Çok sınıflı problemlerde, belirsizlik ölçütü olarak farklı yaklaşımlar önerilmiştir. Örneğin, entropi temelli belirsizlik ölçütü, çoklu sınıflandırma, güven aralığı gibi farklı ölçütler kullanılabilir. Bu ölçütler, sınıflandırıcının belirsizliğini ölçerek, en iyi etiketleme için örnekleri seçmeye yarar. Ayrıca, çok sınıflı problemler için önerilmiş özel ölçütler vardır, örneğin Marjin Örnekleme (MS) yöntemi, Scheffer ve arkadaşları [17] tarafından önerilen bir örnektir. Bu yöntem sınıflandırıcının önerdiği sınıfın marjini kullanarak belirsizliği ölçer.

$$\text{Belirsizlik}(x) = \text{argmax} P_{\theta}(y_1/x) - P_{\theta}(y_2/x) \quad (10)$$

En yüksek olasılığa sahip birinci ve ikinci sınıf etiketlerini gösteren y_1 ve y_2 parametreleri Denklem 10' da gösterilmiştir.

Marjin, sınıflandırıcının önerdiği sınıf ile en yakın olası sınıf arasındaki farktır. Örneklerin marjin değeri yüksekse, sınıflandırıcının kararı daha emindir ve bu örnekler etiketleme için daha az önemlidir. Ayrıca, marjin değeri düşükse, sınıflandırıcının kararı daha belirsizdir ve bu örnekler etiketleme için daha önemlidir. Marjin ölçütü belirsizlik ölçütü olarak kullanılır, yüksek marjin değerine sahip örneklerin belirsizliği düşük, düşük marjin değerine sahip örneklerin belirsizliği yüksektir. Örnekler arasındaki marjin değeri küçüldükçe sınıflandırıcının karar vermesi zorlaşmaktadır. Bu nedenle sınıflandırıcının kararsız kalabileceği örnekler etiketleme için seçilmektedir.

DVM sınıflandırıcılar son yıllarda sınıflandırma problemlerinde oldukça başarılı sonuçlar üretmektedirler. Ayrıca, AÖ uygulamalarında farklı bir yaklaşım olarak da çoğunlukla kullanılmaktadırlar. DVM'nin genelleştirme yeteneği AÖ uygulamalarında yüksek başarı sağlamaktadır. En popüler ve etkili AL yaklaşımı olarak DVM tabanlı Marjin Örnekleme (MÖ) kabul edilir. MÖ'da, belirsizlik ölçütü olarak DVM sınıflandırıcı tarafından belirlenen ayırıcı düzlem uzaklıkları kullanılır. En yakın örnekler ayırıcı düzlem olarak seçilir ve iteratif olarak uzman tarafından etiketlenir.



Şekil 5. MÖ aktif öğrenme.

Bu bölümde sunulan aktif öğrenme yaklaşımlarının amacı, uzman tarafından etiketlenen en belirsiz örneklerle sınıflandırıcının performansını arttırmak olarak açıklanmıştır. Ancak, bu yaklaşımların hızının darboğaz oluşturması problem olarak ortaya çıkmıştır. Bu problem için önerilen çözümler arasında, birden fazla belirsiz örnek içeren örnekler kümesinin seçilmesi yerine, sadece en belirsiz örneğin seçilmesi yer almaktadır. Bu yaklaşım, hız artışı sağlamakla birlikte, benzer örneklerin seçilmesi sınıflandırıcı performansını etkileyebilir. Bu problemi çözmek için çeşitlilik (diversity) ölçütü önerilmiştir [18]. Literatürde, çeşitlilik ölçütü için farklı yöntemler mevcuttur, bunlar arasında kümeleme tabanlı, açı tabanlı ve en yakın destek vektör tabanlı yöntemler yer almaktadır.

Bu çalışmada, aktif öğrenme yaklaşımı ile haploid ve diploid mısır taneleri yüksek doğrulukla sınıflandırılmıştır. Önerilen sınıflandırma sisteminde temel strateji, ayırt edici bilgiler taşıyan minimum sayıda mısır örneği kullanarak yüksek başarılı sonuçlar elde etmektir. Böylece yüksek sınıflandırma başarımı korunarak etiketleme maliyetleri en aza indirgenecektir.

4. DENEYLER VE GÖZLEMLER

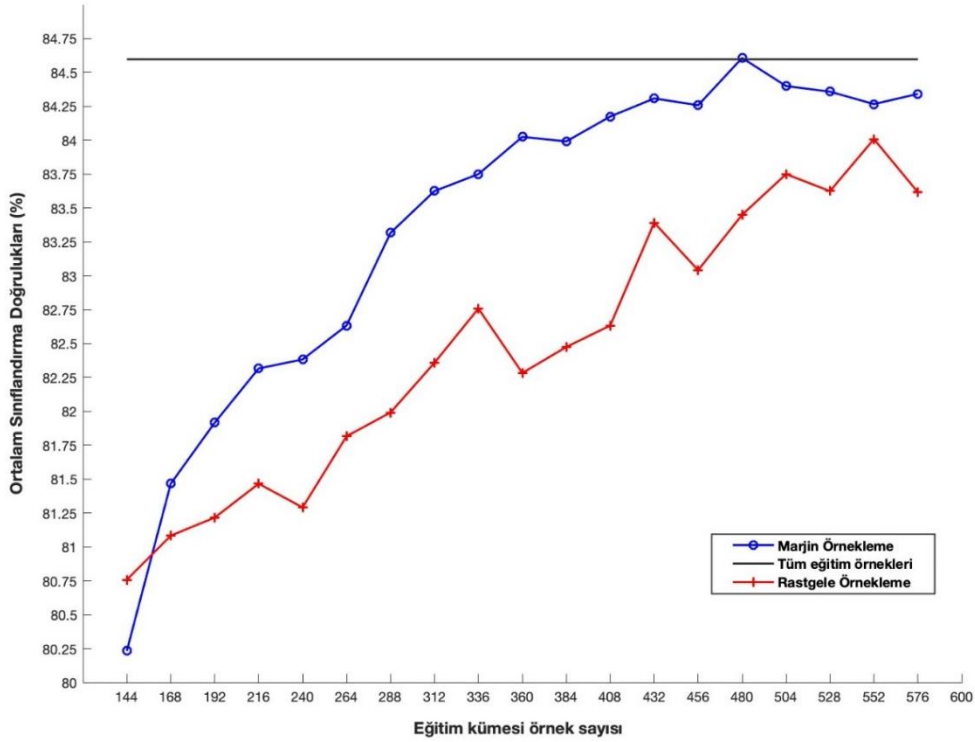
Deneyel çalışmalarda mısır tanesi görüntülerinden elde edilen 12 farklı renk momenti öznelik olarak kullanılmıştır. Sınıflandırma deneyleri için kullanılan eğitim ve test kümeleri 10 farklı deneme ile seçilmiştir. Sınıflandırma çalışmalarında eğitim ve test kümeleri rastgele 10 farklı deneme ile seçilmiş ve kullanılmıştır. 1200 adet diploid, 1200 adet haploid olmak üzere 2400 mısır tanesi kullanılmıştır. Radyal Tabanlı Fonksiyon (RTF) kullanılarak DVM sınıflandırıcısının çekirdek fonksiyonu oluşturulmuştur. İlgili C ve γ parametreleri ise ızgara-arama (grid-search) yöntemi ile tespit edilmiştir. AÖ deneylerinde başlangıç eğitim kümesi T için rastgele seçilen 144 mısır örneği kullanılmış ve DVM sınıflandırıcı eğitilmiştir. Başlangıç eğitim kümesi, örnekler havuzu ve test kümesinde yer alan örnek sayıları Tablo 1'de verilmiştir.

Tablo 1. Eğitim ve test kümelerinden yer alan örnek sayıları.

Mısır Sınıfları	T	P	TS
Haploid	72	528	600
Diploid	72	528	600
Toplam	144	1056	1200

Eğitilen DVM sınıflandırıcı, etiketsiz mısır örnekleri havuzunu sınıflandırmak için kullanılmıştır. Sınıflandırılan örnekler arasından, En İyi Örnekleme (MÖ) stratejisi kullanarak en uygun 24 mısır örneği seçilerek eğitim kümesine eklenmiştir. Bu işlem, iteratif olarak etiketsiz örnekler havuzu (U) boşaltılana kadar devam eder. Her iterasyonda eğitim kümesine yeni örnekler eklendiği için, DVM sınıflandırıcısı her iterasyonda tekrar eğitilir. Bu çalışmada geliştirilen MÖ yönteminin sonuçları, Rastgele Örnekleme (Random Örnekleme (RÖ)) yönteminin sonuçları ile karşılaştırılarak performans testi yapılmıştır. Test yöntemi olarak belirlenen RÖ stratejisinde, her bir AÖ iterasyonu için U etiketsiz örnekler havuzundan rastgele 24 örnek seçilir ve eğitim kümesine etiketlenerek eklenir. Deneyel çalışmalarda elde edilen MÖ, RÖ ve tüm eğitim kümesine ait ortalama sınıflandırma sonuçları grafiksel olarak Şekil 6'da sunulmuştur.

Şekil 6'da yer alan sınıflandırma sonuçlarından görüleceği üzere, eğitim kümesindeki 1200 adet mısır örneğinin tamamı kullanılarak yapılan sınıflandırma işleminde %84.59 başarı oranı elde edilmiştir. Aynı parametreleri kullanarak geleneksel sınıflandırma yaklaşımı (eğitim kümesindeki tüm örneklerin kullanılması) ile karşılaştırmak üzere MÖ ve RÖ aktif öğrenme stratejilerine ait sonuçlarda Şekil 6'da sunulmuştur. Sınıflandırma sonuçlarının objektif olması için her iterasyonda 10 farklı sınıflandırma sonucunun ortalaması alınarak sunulmuştur.



Şekil 6. Eğitim kümesindeki örnek sayısına bağlı ortalama MÖ ve RÖ sınıflandırma başarımları. Düz çizgi ise tüm örneklerin eğitim kümesinde yer alması durumunda elde edilen ortalama sınıflandırma başarısını göstermektedir.

Şekil 6'da yer alan mavi renkli eğriden görüleceği üzere 144 adet mısır örneğinden oluşan başlangıç eğitim kümesi ile yapılan sınıflandırma işlemi, her iterasyonda en çok bilgi taşıyan (ayırıcı düzleme en yakın örnekler) 24 örneğin seçilerek eğitim kümesinin zenginleştirilmesiyle devam etmiştir. Hedef sınıflandırma başarımı olan %84,59 doğruluk oranına 15. iterasyonda 480 örnek kullanılarak ulaşılmıştır. MÖ stratejisine ait sonuçları karşılaştırmak için bu çalışmada Rastgele Örnekleme (RÖ) yöntemi de sınıflandırma süreçlerinde kullanılmıştır. Şekil 6'da kırmızı renkli sınıflandırma sonuçları eğrisi bize MÖ yönteminden farklı olarak örnekler havuzu olan P 'den rastgele olarak seçilen örneklerin eğitim kümesine eklenmesiyle elde edilen sınıflandırma sonuçlarını göstermektedir.

Bu problemde karşılaşılan önemli darboğazlardan birisi örneklerin etiketlenmesi için gereken zaman ve para maliyetidir. Bu çalışmada önerilen Aktif Öğrenme tabanlı sınıflandırma yöntemi ile çok daha az etiketli örnek kullanılarak yüksek sınıflandırma sonuçları elde edilebilmektedir. Geleneksel sınıflandırma yaklaşımlarında 1200 mısır örneği ile bu sınıflandırma işlemi yapılırken, önerilen AÖ yaklaşımli sınıflandırıcı sisteminde seçilen en iyi 480 mısır örneği ile aynı sonuçlar elde edilebilmektedir.

5. SONUÇ VE DEĞERLENDİRME

Mısır, çok yönlü kullanım alanı ve verimliliği ile halen dünya popülasyonunun temel gıda ihtiyacını karşılayan altı tahıldan bir tanesidir. 1800'lü yıllardan günümüze kadar mısır hemen hemen tüm dünyada ıslah çalışmalarının en yoğun şekilde sürdürüldüğü bir bitkidir. Mısır ıslah süreçlerinde çeşitli geleneksel yöntemler kullanılarak örneklerin haploid, diploid sınıflandırmaları yapılabilmektedir. Ancak bu yöntemlerin hata oranlarının yüksekliği, zaman ve para maliyetleri önemli dezavantajlardır. Bu çalışmada, yüksek doğruluklu ve minimum maliyetli bir sınıflandırıcı oluşturulabilmesine imkân tanıyan aktif öğrenmeye yaklaşımı kullanılmıştır. Önerilen çalışmada mısır tanelerine ait en iyi renk moment öznelikleri seçilmiş ve AÖ yaklaşımı ile DVM sınıflandırıcı kullanılarak minimum sayıda eğitim örneği kullanılarak tatmin edici sonuçlar elde edilmiştir. Geleneksel sınıflandırma yaklaşımında 1200 mısır örneği ile sınıflandırıcı eğitilecek iken bu çalışmada aynı sonuçlar 480 mısır örneği kullanılarak elde edilmiştir. Elde edilen sonuçlar DVM sınıflandırıcı için etiketleme maliyetlerini minimuma indirmiştir.

Yazar Katkıları

Yazarlar makaleye eşit derecede katkı sağlamıştır.

Çıkar Çatışması

Makale yazarları aralarında herhangi bir çıkar çatışması olmadığını beyan ederler.

KAYNAKÇA

- [1] B.W. Boote, D.J. Freppon, G.N. de La Fuente, T. Lübberstedt, B.J. Nikolau, and E.A. Smith “Haploid differentiation in maize kernels based on fluorescence imaging”, *Plant Breeding*, vol. 35, no. 4, pp. 439-445 2016.
- [2] G.N. de La Fuente, J.M. Carstensen, M.A. Edberg, and T. Lübberstedt “Discrimination of haploid and diploid maize kernels via multispectral imaging”, *Plant Breeding*, vol. 136, no. 1, pp. 50-60, 2017.
- [3] J. Lin, L. Yu, W. Li, and H. Qin “Method for Identifying Maize Haploid Seeds by Applying Diffuse Transmission Near-Infrared Spectroscopy”, *Applied Spectroscopy*, vol. 72, no. 4, pp. 611-617, 2018.
- [4] Y. Altuntas, A.F. Kocamaz, R. Cengiz, and M. Esmeray “Classification of haploid and diploid maize seeds by using image processing techniques and support vector machines”, *26th IEEE Signal Processing and Communications Applications Conference*, 2018.
- [5] X.-Y. Wang, W.-X. Liao, D. An, and Y. Wei “Maize Haploid Identification via LSTM-CNN and Hyperspectral Imaging Technology”, *CoRR arXiv*, vol. abs/1805.0, 2018.
- [6] L. Yu, W. Liu, W. Li, H. Qin, J. Xu, and M. Zuo “Non-destructive identification of maize haploid seeds using nonlinear analysis method based on their near-infrared spectra”, *Biosystem Engineering*, vol. 172, pp. 144-153, 2018.
- [7] Y. Altuntaş, Z. Cömert, and A.F. Kocamaz “Identification of haploid and diploid maize seeds using convolutional neural networks and a transfer learning approach”, *Comp. and Elect. in Agr.*, vol. 163, pp. 1-11, 2019.
- [8] E. Donmez “Discrimination of Haploid and Diploid Maize Seeds Based on Deep Features”, *28th Signal Processing and Communications Applications Conference*, 2020.
- [9] E. Dönmez “Classification of Haploid and Diploid Maize Seeds based on Pre-Trained Convolutional Neural Networks”, *Celal Bayar Üniversitesi Fen Bilimleri Dergisi*, vol. 16, no. 3, pp. 323-331, 2020.
- [10] E. Dönmez “Enhancing classification capacity of CNN models with deep feature selection and fusion: A case study on maize seed classification”, *Data and Know. Eng.*, vol. 141, p. 102075, pp. 1-12, 2022.
- [11] A. Gunes, A.S. Bilgi, G. Ortac, H. Kalkan, and K. Tasdemir “Active learning method for classifying the mold contaminated figs”, *24th Signal Processing and Communication Application Conference*, 2016.
- [12] H. Kalkan, A. Güneş, E. Durmuş, and A. Kuşçu “Non-invasive detection of aflatoxin-contaminated figs using fluorescence and multispectral imaging”, *Food Additives and Contaminants - Part A*, vol. 31, no. 8, pp. 1414-1421, 2014.
- [13] P. Pudil, J. Novovičová, and J. Kittler “Floating search methods in feature selection”, *Pattern Recog. Lett.*, vol. 15, no. 11, pp. 1119-1125, 1994.
- [14] D.D. Lewis and W.A. Gale “A sequential algorithm for training text classifiers”, *17th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 1994.
- [15] H.S. Seung, M. Opper, and H. Sompolinsky “Query by committee”, *Fifth Annual ACM Workshop on Computational Learning Theory*, 1992.
- [16] C.E. Shannon “A Mathematical Theory of Communication”, *Bell System Technical Journal*, vol. 27, no. 3, pp. 379-423, 1948.
- [17] T. Scheffer, C. Decomain, and S. Wrobel “Active hidden markov models for information extraction”, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 2189, 2001.
- [18] K. Blinker “Incorporating Diversity in Active Learning with Support Vector Machines”, *Twentieth International Conference on Machine Learning*, 2003.