



# Kahramanmaraş Sütçü İmam University

## Journal of Engineering Sciences



Geliş Tarihi : 26.05.2024  
Kabul Tarihi : 05.10.2024

Received Date : 26.05.2024  
Accepted Date : 05.10.2024

### FARKLI ÇAPRAZLAMA TEKNİKLERİ KULLANAN DAĞINIK ARAMA ALGORİTMASI İLE EVRİŞİMLİ SİNİR AĞLARINDA HİPER PARAMETRE OPTİMİZASYONU

### HYPERPARAMETER OPTIMIZATION IN CONVOLUTIONAL NEURAL NETWORKS WITH SCATTER SEARCH ALGORITHM USING DIFFERENT CROSSOVER TECHNIQUES

Hediye ORHAN<sup>1\*</sup> (ORCID: 0000-0001-8760-914X)  
Dilara Sevim POLAT<sup>1</sup> (ORCID: 0009-0006-6985-1761)  
Hüseyin Haklı<sup>1</sup> (ORCID: 0000-0001-5019-071X)

<sup>1</sup> Necmettin Erbakan Üniversitesi, Bilgisayar Mühendisliği Bölümü, Konya, Türkiye

\*Sorumlu Yazar / Corresponding Author: Hediye ORHAN, hediyeorhan2015@gmail.com

#### ÖZET

Günümüzde yapay zekâ uygulamaları hayatın her alanında kullanılmaktadır ve gün geçtikçe daha uygulanabilir öneriler ve sonuçlar sunar hale gelmiştir. Evrişimli Sinir Ağları (ESA), birçok gerçek dünya probleminde başarılı sonuçlar veren, son yıllarda etkili ve yoğun bir şekilde uygulanan yapay zekâ algoritmalarından biridir. ESA'lar genellikle görsel bilginin analiz edilmesinde kullanılmaktadır. Görsel bilgi, ESA'larda bulunan evrişim, aktivasyon, havuzlama ve tam bağlantılı katmanlardan geçirilerek analiz edilmektedir. ESA eğitiminde kullanılan veri setine ve karşılaşılan probleme göre çeşitli parametreler kullanılmaktadır. Bu çalışmada en yüksek doğruluk değerini veren hiper parametrelerin seçilebilmesi için ESA eğitiminde kullanılan parametreler ve ağ yapısının oluşturulmasında kullanılan katmanlar optimize edilmiştir. Hiper parametrelerden kanal sayısı, evrişimsel katman, minimum parti boyutu ve aktivasyon fonksiyonu için ayrıık değerler, öğrenme oranı için sürekli değerler belirlenmiştir. Bu çalışmada hiper parametre optimizasyonunu gerçekleştirmek için Dağınık Arama (DA) ve Genetik Algoritmalar (GA) yöntemleri tercih edilmiştir. DA yöntemi, GA ile kıyaslandığında ESA'lar için uygun hiper parametre değerlerinin kolaylıkla belirlenmesini sağlamıştır ve daha yüksek doğruluk değeri elde edilmiştir. Çalışmada elde edilen en yüksek doğruluk değerleri GA yöntemi ile %88.76 iken DA yöntemi ile % 93.24'tür. Bu değer 16 kanal sayısı, 5 x 5 evrişimsel katman, 64 minimum parti boyutu, 0.0052 öğrenme oranı ve reluLayer aktivasyon fonksiyonu parametreleri ile elde edilmiştir.

**Anahtar Kelimeler:** Optimizasyon, dağınık arama, derin öğrenme, hiper parametre

#### ABSTRACT

Nowadays, artificial intelligence applications are used in all areas of life and have become more and more applicable to provide recommendations and results. Convolutional Neural Networks (CNN) is one of the most effective and intensively applied artificial intelligence algorithms in recent years, providing successful results in many real-world problems. CNNs are generally used to analyze visual information. Visual information is analyzed by passing it through convolution, activation, pooling and fully connected layers in CNNs. In CNN training, various parameters are used according to the data set used and the problem encountered. However, finding the best hyperparameter values for a CNN is still a challenging task. In this study, the parameters used in CNN training and the layers used in the network structure are optimized in order to easily select the hyperparameters that give the highest accuracy. Discrete values for the number of channels, convolutional layer, minimum batch size and activation function and continuous values for the learning rate were chosen as hyper-parameters. In this study, Scatter Search (SS) algorithm is preferred to perform hyper parameter optimization. With the SS method, appropriate hyperparameter values for

ToCite: ORHAN, H., & POLAT, D.S., & HAKLI, H., (2024). FARKLI ÇAPRAZLAMA TEKNİKLERİ KULLANAN DAĞINIK ARAMA ALGORİTMASI İLE EVRİŞİMLİ SİNİR AĞLARINDA HİPER PARAMETRE OPTİMİZASYONU. *Kahramanmaraş Sütçü İmam Üniversitesi Mühendislik Bilimleri Dergisi*, 27(4), 1437-1450.

CNNs were easily determined and a higher accuracy was achieved. The highest accuracy value obtained in the study is 93.24%. This value was obtained with 16 number of channels, 5 x 5 convolutional layers, 64 minimum batch size, 0.0052 learning rate and reluLayer activation function parameters.

**Keywords:** Optimization, scatter search, deep learning, hyperparameter

## GİRİŞ

Evrişimli Sinir Ağları (ESA), görüntü işleme, ses tanıma, doğal dil işleme ve diğer birçok yapay zekâ uygulamasında başarıyla kullanılan derin öğrenme modelidir. ESA'lar, özellikle görsel verilerin analizinde oldukça başarılıdır. ESA'lar kullandığı katmanlar aracılığı ile veriler arasındaki ilişkileri öğrenmektedir ve özelliklerin hiyerarşik temsillerini oluşturmaktadır. ESA'larda bulunan katmanlar, bu katmanların parametre değerleri, katman sayısı, katmanlarda bulunan filtre sayısı ve filtre boyutu gibi birçok kriter elde edilen doğruluğu etkilemektedir. Böylece doğru hiper parametre seçimi ile daha iyi doğruluk oranları elde edilebilmektedir (Bozkurt & Yağanoğlu, 2021). ESA, yapay zekâ alanında önemli bir yere sahiptir ve günümüzde birçok uygulamada etkili bir şekilde kullanılmaktadırlar. Hiper parametrelerin optimizasyon algoritmaları ile belirlenmesi, birçok ihtimalin bir arada bulunmasını sağlamaktadır (Zhang vd., 2021). Bu sonuçları gözlemleyen kullanıcı en yüksek doğruluk değerini veren parametreleri kolaylıkla seçebilmektedir.

Xiao vd. (2020) CIFAR-10 adlı veri seti üzerinde hiper parametre optimizasyonu gerçekleştirmiştir. Çalışmada, değişken uzunluklu GA kullanılarak yapılan hiper parametre optimizasyonunda elde edilen en iyi doğruluk değeri %88.92'dir. En iyi sonuç toplamda yaklaşık 25 saat süren bir optimizasyon süreci sonucunda elde edilmiştir. Buna ek olarak çalışmada rastgele arama (RA), klasik GA ve büyük ölçekli evrim algoritması ile karşılaştırmalar yapılmış ve sınırlı GPU kaynaklarına sahip durumlarda değişken uzunluklu GA'nın daha iyi sonuçlar verdiği belirtilmiştir. Çalışmada optimize edilen parametrelerin dikkate alınmasıyla, ele alınacak parametreler hakkında daha iyi bir anlayış elde edilmiştir.

Tanyıldızı vd. (2019) çalışmasında hiper parametre optimizasyonu için ızgara arama yöntemi kullanmıştır. Fakat bu yöntemin çok fazla işlem gücü ve zaman gerektirmesi nedeniyle hızını arttırmak amacıyla parçacık sürü optimizasyonu (PSO) algoritması eklenmiştir.

Özbay vd. (2023) çalışmasında ISIC-2019 veri seti ile cilt kanseri sınıflandırılması yapılmıştır. Sınıflandırma için oluşturulan ESA hiper parametreleri PSO algoritması ile optimize edilmiştir. 50 devir ile bir optimizasyon gerçekleştirilmiştir. Kullanılan PSO algoritması ile %99.33 doğruluk değeri elde edilmiştir. Elde edilen sonuçlardan PSO algoritmasının GA, diferansiyel evrim (DE) algoritması ve gri kurt optimizasyonu (GKO) algoritmasından daha iyi sonuç verdiği gözlemlenmiştir.

Fujino vd. (2017) el çizimleri ile ilgili bir veri setinin sınıflandırılması çalışmasında AlexNet mimarisi için hiper parametre optimizasyonu GA ile gerçekleştirilmiştir. MNIST veri seti için GA ve PSO algoritmasında hiper parametre optimizasyonları gerçekleştirilmiştir.

GA ile Bochinski vd. (2017) çalışmasında MNIST veri seti kullanılarak evrimsel algoritma ile %99.76 doğruluk değeri elde edilmiştir. Çalışmada aktivasyon fonksiyonu olarak Relu, öğrenme oranı olarak 0.0001 değeri ve optimizer olarak Adam algoritması tercih edilmiştir. %50 seyreltme oranı uygulanmıştır. Eğitim 30 devirde ve 100 minimum parti boyutunda gerçekleştirilmiştir. Optimize edilen hiper parametrelerden evrişimsel katmanlar, çekirdek boyutu ve tam bağlantılı katmanlar ise belirli aralıklarda değerler almıştır.

Lorenzo vd. (2017) çalışmasında, PSO algoritması ile %99.45 doğruluk değeri elde etmiştir. Çalışmada çoklu sınıflandırma üzerine odaklanılmıştır. Bu nedenle MNIST ve CIFAR-10 veri setleri üzerinde çalışılmıştır.

İnik (2023) ESC-10, ESC-50 (Piczak, 2015) ve Urbansound8k (Salamon, Jacoby, & Bello, 2014) veri setleri üzerinde PSO algoritması ile bir ESA hiper parametre optimizasyonu gerçekleştirmiştir. ESC-10 veri setinde %98.64, ESC-50 veri setinde %93.71 ve Urbansound8k veri setinde %98.45 başarı oranı elde edilmiştir.

Andonie ve Florea (2020) CIFAR-10 veri seti üzerinde hiper parametre optimizasyonu için RA, nelder-mead, PSO, sobol dizileri, bayes optimizasyonu, ağaç yapılı parzen tahmini ve ağırlıklı RA algoritmalarını test etmiştir. Çalışmada en yüksek doğruluk değeri %85 ile ARA algoritmasından elde edilmiştir.

Raziani ve Azimbagirad (2022) insan aktiviteleri ile ilgili HAR veri seti (Anguita, Ghio, Oneto, Parra, & Reyes-Ortiz, 2013) üzerinde yedi farklı algoritma ile hiper parametre optimizasyonunu test etmiştir. Bu algoritmalar; Güve Alev Optimizasyonu (GAO), GKO, Çok Evren Optimizasyonu (ÇEO), Sinüs Kosinüs Algoritması (SKA), Salp Sürü Algoritması (SSA), Balina Optimizasyon Algoritması (BOA) ve PSO algoritmasıdır. Çalışmada havuzlama katmanı boyutu, parti boyutu, çekirdek boyutu, devir ve filtre sayıları optimize edilmiştir. Algoritmaların 50 devir çalışması sonucunda elde edilen kayıp fonksiyonu grafiğine bakıldığında en başarılı sonucun GKO algoritması ile elde edildiği görülmektedir.

Gülcü ve Kuş (2019) PSO, BO, GA ve DE algoritmaları ile ESA hiper parametre optimizasyonu gerçekleştirmiştir. Çalışmada filtre sayısı, filtre boyutu, katman sayısı, seyreltme oranı, öğrenme oranı ve parti boyutu parametreleri optimize edilmiştir. En başarılı sonuç GA ve PSO algoritmalarından elde edilmiştir.

Kıymaç'ın (2022) çalışmasında, EKG sinyallerinin sınıflandırılması için açık erişimli veri tabanından bir veri seti kullanılmıştır. Veritabanı, 47 konudan elde edilen, örnekleme hızı 360 Hz olan 48 yarım saatlik EKG kaydı içermektedir. Çalışmada aktivasyon fonksiyonu olarak gizli katmanlarda Relu, çıkış katmanında softmax kullanılmıştır. Optimizer olarak Adam algoritması tercih edilmiştir. Oluşturulan yapay sinir ağı 10 devir ile eğitilmiştir. Optimize edilecek parti boyutu, öğrenme oranı, evrimsel katman ve seyreltme oranı gibi hiper parametreler için farklı ayırık değerler belirlenmiştir. Çalışmada en iyi doğruluk değeri PSO ile elde %98.87 olarak elde edilmiştir.

Yurdakul'un (2022) çalışmasında, filtre sayısı ve boyutu, aktivasyon fonksiyonu, havuzlama algoritması, tam bağlantılı katman sayısı, nöron sayısı ve öğrenme oranı farklı meta-sezgisel algoritmalar ile belirlenmiştir. Çalışmada CIFAR-10 veri seti üzerinde en yüksek doğruluk oranı yapay arı kolonisi algoritması ile %87.74 olarak elde edilmiştir. Çalışma sonucunda hiper parametreler için optimum değer aralıkları; filtre boyutu için 5 ile 9 arası, filtre sayısı için 64 ve 256 arası, nöron sayısı için 32 ve 256 arası, seyreltme oranı için 0.5 ve 0.7 arasındadır. Aktivasyon fonksiyonu için optimum fonksiyonun Relu olduğu belirtilmiştir.

Atteia vd. (2022) çalışmasında diyabetik makulopati tanısında kullanılan ESA hiper parametreleri optimize edilmiştir. Çalışmada, kamuya açık retina görüntü veri seti kullanılmıştır. Bu veri setleri, fundus fotoğrafçılığı ve optik koherens tomografi olmak üzere iki tarama tekniği ile yakalanan retina görüntülerini içermektedir. Fundus retinografi görüntü seti, Indian Diabetic Retinopathy Image Dataset (IDRiD) (Porwal vd., 2018) olarak adlandırılmaktadır. Bu çalışmada kullanılan optik koherans tomografi retina görüntüleri veri seti ise Kaggle Dataset'dir (Mooney, 2018). Çalışmada, ESA hiper parametrelerinin optimizasyonu için Bayesian optimizasyon yöntemi kullanılmıştır. Bu yöntem, modelin performansını artırmak amacıyla hiper parametrelerin en uygun değerlerini bulmak için bir akış fonksiyonu kullanarak arama alanını örnekler. Sonuçlar, Bayesian optimizasyonu ile elde edilen ESA modelinin, optimize edilmemiş bir ağa göre daha iyi performans gösterdiğini ortaya koymaktadır. Öğrenme oranı, katmanlardaki nöron sayısı, aktivasyon fonksiyonu ve mini-batch boyutu gibi hiper parametreler optimize edilmiştir. Çalışmada, optimize edilmiş model test setinde %94.2 doğruluk oranı elde etmiştir. Bu oran, optimize edilmemiş modelin %93.2 doğruluk oranına göre belirgin bir iyileşme göstermektedir. Modelin performansı doğruluk, hassasiyet ve kesinlik metrikleri ile değerlendirilmiştir.

Amou vd. (2022) çalışmasında beyin tümörünün ESA ile sınıflandırılması gerçekleştirilmiştir. Bayesian optimizasyonu ile hiper parametreler optimize edilmiştir. Çalışmada optimize edilen hiper parametreler şu şekildedir; aktivasyon fonksiyonları olarak ReLU, ELU, Sigmoid, SELU ve Tanh belirlenmiştir. Minimum parti boyutu 1 ile 128 arasında uygun değerler almaktadır. Seyreltme oranı %0.1 ile %0.5 arasında değerler almaktadır. Yoğun katman düğüm sayısı 32 ile 1024 arasında değerler almaktadır. Gradyan iniş optimizasyonu için Adam, Nadam, AdaMax, RMSProp ve SGD algoritmaları belirlenmiştir. Bayesian optimizasyonu ile bu parametrelerden en başarılı sonucu veren hiper parametre kombinasyonu elde edilmiştir. Optimize edilen ESA modelinin %98.70 doğrulama doğruluğuna ulaştığı belirtilmiştir. Buna karşın, VGG16, VGG19, ResNet50, InceptionV3 ve DenseNet201 modellerinden sırasıyla %97.08, %96.43, %89.29, %92.86 ve %94.81 doğrulama doğruluğu elde edildiği belirtilmiştir. Doğruluk ve zaman bakımından en başarılı sonuç Tanh aktivasyon fonksiyonu, 1 minimum parti

boyutu, 0.25 seyreltme oranı, 32 yoğun katman ve SGD optimizasyon algoritması ile elde edilmiştir. Toplamda 30 devirlik bir çalışma süresinin 11'inci devrinde sonuca ulaşılmıştır.

ZainEldin vd. (2022) çalışmasında beyin tümörü sınıflandırılması gerçekleştirilmiştir. Çalışmada Uyarlanabilir Dinamik Sinüs Cosine Uygunluk Gri Kurt algoritması kullanılmıştır. Bu algoritma hem sine cosine hem de gri kurt algoritmalarının güçlü yönlerini bir araya getirerek esnek bir çerçeve sunmaktadır. Çalışmada; öğrenme oranı, momentum oranı, minimum parti boyutu, L2 düzenleme oranı ve devir sayısı gibi birçok parametre optimize edilmiştir. Hiper parametre optimizasyonu sonucunda %99.98 doğruluk değeri elde edilmiştir. Model değerlendirilmesinde doğruluk, kesinlik, hassasiyet ve F1-skor performans metrikleri kullanılmıştır.

Aslan vd. (2022) çalışmasında COVID-19 teşhisi için bilgisayarlı tomografi göğüs görüntülerinin sınıflandırılması üzerine odaklanılmıştır. Çalışmada, sekiz farklı popüler ESA modeli (AlexNet, ResNet18, ResNet50, InceptionV3, DenseNet201, InceptionResNetV2, GoogleNet, MobileNetV2) kullanılarak özellikler çıkarılmıştır. Elde edilen özellikler, dört farklı makine öğrenimi yöntemi ile sınıflandırılmıştır. Bunlar: Destek Vektör Makineleri, k-En Yakın Komşu, Naive Bayes ve Karar Ağaçlarıdır. Her makine öğrenimi algoritması için hiper parametrelerin optimize edilmesi amacıyla Bayesian optimizasyon yöntemi kullanılmıştır. Çalışmada, optimizasyon sayesinde elde edilen en yüksek sınıflandırma doğruluğu %96.29 olarak rapor edilmiştir. DenseNet201 modeli ile SVM algoritmasının birlikte kullanılması ile bu başarıya ulaşılmıştır. Çalışmada optimize edilen hiper parametreler kullanılan makine öğrenimi algoritmasına göre farklılık göstermektedir. K-En Yakın Komşu algoritmasında mesafe ölçütü ve komşu sayısı, Destek Vektör Makineleri algoritmasında kutup kısıtları, kodlama yöntemi ve çekirdek ölçeği, Karar Ağacı algoritmasında minimum yaprak boyutu ve Naive Bayes algoritmasında çekirdek genişliği ve dağılım türü optimize edilmiştir.

Sethi vd. (2021) çalışmasında Alzheimer hastalığının erken teşhisi için çalışılmıştır. Çalışmada veri artırma yöntemleri uygulanmıştır. Derin öğrenme algoritması olarak ESA ve Uzun Kısa Süreli Bellek (LSTM) kullanılmıştır. Çalışmada Bayesian optimizasyon yöntemi kullanılmıştır. Bayesian optimizasyon, önceki deneyimlerden elde edilen sonuçları kullanarak yeni hiper parametre kombinasyonlarının seçimini yönlendirmektedir, böylece daha az deneme ile daha iyi sonuçlar elde edilmesine yardımcı olmaktadır. Öğrenme oranı, optimizasyon algoritmaları, parti boyutu, aktivasyon fonksiyonları ve düzenleme parametreleri (L1, L2) gibi hiper parametreler optimize edilmiştir. Elde edilen doğruluk değeri 0.0020 öğrenme oranı ve Adamax optimizasyon algoritması ile %86.28'dir. Çalışmada verilere artırma yöntemi uygulanarak LSTM ve ESA modelinin hibrit kullanımından %92.56 doğruluk değeri elde edilmiştir.

Zhang vd. (2021) çalışmasında ESA tabanlı akciğer nodülü sınıflandırmasını araştırmaktadır. Çalışmada hiper parametre optimizasyonu için bir Gaussian surrogat modeli kullanılmıştır. Bu model, hiper parametre konfigürasyonlarının hata fonksiyonunu yaklaşık olarak tahmin etmek için tasarlanmıştır. Ayrıca, hiper parametre optimizasyonu sürecinde, bir evrimsel strateji ile birlikte hiper parametrelerin önemine dayalı bir arama yöntemi uygulanmıştır. Çalışmada öğrenme oranı, parti boyutu ve devir sayısı gibi parametreler optimize edilmiştir. Çoklu görünüm ESA kullanılarak yapılan çalışmalarda, ikili sınıflandırma için %94.59 ve üçlü sınıflandırma için %86.09 doğruluk değerleri elde edilmiştir. Çok ölçekli ESA yapısı, ikili sınıflandırmada %92.21 ve üçlü sınıflandırmada %84.81 doğruluk sağlamıştır. Çoklu görünüm ESA'da amaç, çoklu görüntü kaynaklarından gelen bilgiyi birleştirerek daha iyi sınıflandırma sonuçları elde etmektir. Bu, özellikle nodülün farklı açılardan görünümünü değerlendirmek için faydalıdır. Çok ölçekli ESA'da amaç, farklı ölçeklerdeki özellikleri yakalayıp nodülün boyutuna ve şekline göre daha iyi bir sınıflandırma yapmaktır. Bu, nodüllerin değişken boyutları ve morfolojileri için daha esnek bir yaklaşım sunar.

Shankar vd. (2020) çalışmasında diyabetik retinopati tespiti gerçekleştirilmiştir. Diyabetik retinopati, dünya genelinde artan görme kaybının başlıca nedenlerinden biridir. Hiper parametre optimizasyonu için bu çalışmada Bayesian optimizasyon algoritması kullanılmıştır. Bu yöntem, hedef fonksiyonun optimizasyonunu kolaylaştırmak için bir "surrogate" model kullanmaktadır. Optimize edilen parametreler arasında devir sayısı 500 olarak belirlenmiştir. Öğrenme oranı 0.001 olarak ayarlanmış ve momentum değeri ise 0.9 olarak optimize edilmiştir. Bu parametreler, modelin performansını en üst düzeye çıkarmak için en uygun değerler olarak belirlenmiştir. Çalışmada, etkili bir DR hastalığı teşhis modeli olan HPTI-v4 modeli sunulmuştur. Sunulan HPTI-v4 modeli, histogram ve Inception v4 tabanlı özellik çıkarım süreçlerini içeren segmentasyon sürecini kapsamaktadır. Son olarak, sınıflandırma işlemleri çok katmanlı algılayıcı kullanılarak gerçekleştirilmiştir. Çalışmada sunulan HPTI-v4 modelinden %99.49 maksimum doğruluk, %98.83 hassasiyet ve %99.68 özgüllük değeri elde edilmiştir.



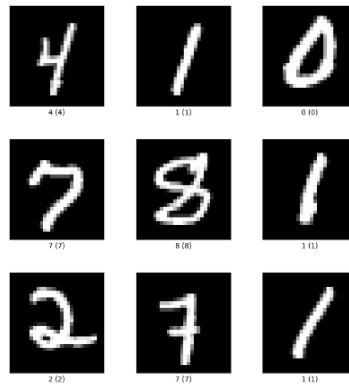
Literatürde hiper parametre optimizasyonunda daha çok PSO, GA ve Bayesian algoritmalarının tercih edildiği görülmüştür. Bu çalışmada etkili ve sistematik bir arama stratejisine sahip olan DA yöntemi hiper parametre optimizasyonu için uygulanmıştır. Çalışmada üç farklı çaprazlama tekniği kullanılarak birçok farklı aday çözüm elde edilmiştir.

Çalışmanın Materyal ve Metod bölümünde, projenin geliştirme aşamaları, kullanılan veri seti hakkında bilgiler verilmiştir. Oluşturulan evrişimsel sinir ağı, optimizasyon için kullanılan DA algoritması ve çaprazlama tekniklerinin çalışma mantığı detaylı olarak açıklanmıştır. Bulgular ve Deneysel Sonuçlar bölümünde, önerilen yöntem açıklanmıştır, elde edilen deneysel sonuçlar grafikler ve görseller ile sunulmuştur. Sonuç bölümünde, kullanılmasına karar verilen algoritma ile ilgili bilgiler ve çalışmanın sonuçlandırılması hakkında özetleyici bilgiler yer almaktadır.

## MATERYAL VE METOD

### Veri Seti

Çalışmada ESA yapısının hiper parametrelerinin test edilmesi için bir veri seti gereklidir. Veri seti üzerinde hiper parametreler ile oluşturulan ESA yapısı eğitilerek bir doğruluk değeri elde edilmektedir. Çalışmada MNIST veri seti kullanılarak bir ESA yapısının hiper parametreleri optimize edilmiştir (MNIST). MNIST veri seti, 28 x 28 piksel boyutunda siyah-beyaz 70.000 el yazısı rakam görüntüsünden oluşmaktadır. Bu görüntüler, 0 ile 9 arasındaki rakamları temsil etmektedir. MNIST veri kümesi, her bir rakam sınıfında farklı el yazısı stillerini içermektedir. Yani aynı rakamın farklı kişiler tarafından farklı şekillerde yazılmış örnekleri bulunmaktadır. Bu, el yazısı tanıma algoritmalarının farklı yazım stillerini tanımayı öğrenmelerine yardımcı olmaktadır. Kullanılan MNIST veri setinden örnek bir kısım Şekil 1'de görülmektedir.



Şekil 1. MNIST Veri Setinden Bir Örnek (MNIST)

### Evrişimli Sinir Ağları Genel Yapısı

ESA yapısı girdi olarak veri setinden bir görüntü almaktadır. Görüntü, belirli değerler taşıyan piksellerden oluşan matrislerdir. Görüntü; evrişim, havuzlama ve tam bağlantılı katmanlardan geçerek eğitilmektedir. Burada evrişim katmanı; görüntüyü ele alan ilk katmandır (Sakib, Ahmed, Kabir, & Ahmed, 2019). Bu katmanda orijinal görsel boyutundan daha küçük bir filtre görsel üzerinde gezmektedir ve bu görsellerden belirli özellikleri yakalamaya çalışmaktadır. Bir görselin matris haline getirilmesi ve evrişim işlemi uygulanması Şekil 2'de görülmektedir. 3x3'lük bir filtre uygulanarak bir çıktı matrisi elde edilmektedir.

Görüntü		Filtre		Özellik Haritası																																											
<table border="1"><tr><td>1</td><td>1</td><td>1</td><td>0</td><td>0</td></tr><tr><td>0</td><td>1</td><td>1</td><td>1</td><td>0</td></tr><tr><td>0</td><td>0</td><td>1</td><td>1</td><td>1</td></tr><tr><td>0</td><td>0</td><td>1</td><td>1</td><td>0</td></tr><tr><td>0</td><td>1</td><td>1</td><td>0</td><td>0</td></tr></table>	1	1	1	0	0	0	1	1	1	0	0	0	1	1	1	0	0	1	1	0	0	1	1	0	0	*	<table border="1"><tr><td>1</td><td>0</td><td>1</td></tr><tr><td>0</td><td>1</td><td>0</td></tr><tr><td>1</td><td>0</td><td>1</td></tr></table>	1	0	1	0	1	0	1	0	1	=	<table border="1"><tr><td>4</td><td>3</td><td>4</td></tr><tr><td>2</td><td>4</td><td>3</td></tr><tr><td>2</td><td>3</td><td>4</td></tr></table>	4	3	4	2	4	3	2	3	4
1	1	1	0	0																																											
0	1	1	1	0																																											
0	0	1	1	1																																											
0	0	1	1	0																																											
0	1	1	0	0																																											
1	0	1																																													
0	1	0																																													
1	0	1																																													
4	3	4																																													
2	4	3																																													
2	3	4																																													

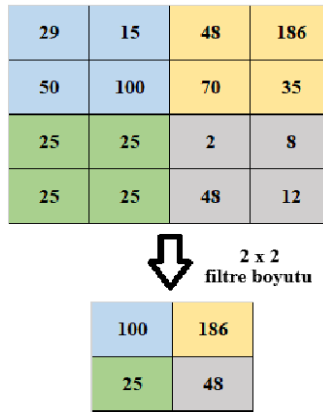
Şekil 2. 3x3'lük Filtre ile Evrişim İşleminin Gerçekleştirilmesi

Filtre uygulanması sonucunda oluşturulan matrise özellik haritası denilmektedir. ESA'da öğrenilen parametreler bu matrislerdeki değerlerdir. Model sürekli bu değerleri güncellemektedir ve özellikleri daha iyi öğrenmeye başlamaktadır. Bu filtreler; köşe bulma, kenar bulma ve bulanıklaştırma gibi özellik filtreleridir. Kenar bulma, giriş bilgisinin yüksek frekanslı bölgelerini temsil etmektedir. Sobel, prewitt, gabor gibi kenar bulma filtreleri mevcuttur. Elde edilen çıkış, görüntülerin kenar bilgisini vermektedir. Kenarlar genelde ESA katmanının en başında hesaplanmaktadır. Filtrelerin girdi üzerinde kaç piksel boyunca kayacağı adım (stride) sayısı ile belirlemektedir. Buna göre özellik haritası boyutu artabilmektedir veya azalabilmektedir.

Kenar bulma, bulanıklaştırma gibi filtreler kullanılırken girdi verisinin boyutu korunmak istenirse dolgulama (padding) işlemi uygulanmaktadır. Dolgulama işlemi çıktı boyutunu korur. Dolgulama işlemi girdi verisinin etrafını 0 ile doldurabilmektedir ya da girdi verisini kopyalayarak çoğaltabilmektedir. Burada temel amaç, bilgi kaybını minimuma indirmektedir.

Aktivasyon fonksiyonu olarak ara katmanlarda genel olarak Relu, Elu, Tanh kullanılmaktadır (Hao, Wang, Liu, & Sun, 2020). Çıktı katmanında sigmoid veya softmax tercih edilmektedir. Sigmoid, [0,1] arasında değerler almaktadır. Türevlenebilir, doğrusal olmayan bir fonksiyondur ve genelde ikili sınıflandırma problemlerinde kullanılmaktadır. Softmax, çoklu sigmoid olarak bilinmektedir ve ikiden fazla sınıf barındıran problemlerde kullanılmaktadır.

Havuzlama katmanında amaç, boyutsallığı azaltmaktır; hem işlem gücü azalır hem de gereksiz özellikler yok sayılarak önemli özelliklere odaklanılmaktadır. Ağın bu katmanında öğrenilen parametre yoktur. Havuzlama katmanında iki tür havuzlama işlemi vardır bunlar; ortalama havuzlama ve maksimum havuzlamadır. Projede kullanılan havuzlama katmanı maksimum havuzlamadır bu işlemde filtre uygulanan kısımdaki en büyük değerler seçilmektedir. Şekil 3'te maksimum havuzlama işlemi görülmektedir.



Şekil 3. Havuzlama Katmanında İşlemlerin Gerçekleştirilmesi

Tam bağlantılı (fully connected) katman genelde ESA mimarisinin sonunda bulunmaktadır. Sınıf skorları gibi hedefleri optimize etmek için kullanılmaktadır. Bu katmanda evrişim işlemleri ve havuzlama işlemleri sonucunda oluşan matris düz vektör haline getirilmektedir. Bu katman, her girişin tüm nöronlara bağlı olduğu bir giriş üzerinde çalışmaktadır.

Tablo 1'de ESA için optimize edilen hiper parametreler görülmektedir. Burada görülmekte olan kanal sayısı parametresi, bir evrişimli katmanda kullanılan filtre sayısını temsil etmektedir. Bu filtreler, giriş verilerindeki farklı özelliklerin çıkarılmasından sorumludur, böylece her bir kanal farklı bir özelliği temsil etmektedir. Kanal sayısı parametresinin artması durumun daha fazla özellik çıkarılabilmektedir, modelin öğrenme kapasitesi artmaktadır. Ancak, bu durum hesaplama maliyetini ve bellek kullanımını artırmaktadır. Azaldığı durumlarda ise çıkarılabilecek özelliklerin sayısı azalmaktadır, bu da modelin performansını olumsuz etkileyebilmektedir. Ancak, hesaplama maliyeti ve bellek kullanımı azalmaktadır. Kanal sayısı parametresine probleme, kullanılan veri setine ve donanıma göre karar verilmektedir.

Evrişimsel katman giriş verisindeki özelliklerin çıkarılmasından sorumludur. Evrişimsel katmanlar filtre boyutunu belirlemektedir. Bu filtreler, giriş verisi üzerinde kayan pencereler gibi çalışarak, belirli boyutlarda yerel bölgelerde özellikler çıkarmaktadır. Filtre boyutunun artması ile daha büyük yerel bölgeler işlenmektedir, bu da daha geniş ölçekli özelliklerin çıkarılmasını sağlamaktadır. Ancak, büyük filtre boyutları hesaplama maliyetini artırmaktadır ve daha fazla bellek gerektirebilmektedir. Filtre boyutunun azalması ile daha küçük yerel bölgeler işlenmektedir, bu da

daha ince detayların çıkarılmasını sağlamaktadır. Küçük filtre boyutları daha az hesaplama maliyeti gerektirmektedir ve daha az bellek kullanmaktadır, ancak çok küçük filtreler yeterince geniş ölçekli özellikleri çıkarmayabilmektedir. Minimum parti boyutu parametresi ise modelin her eğitim veya tahmin aşamasında işlediği en az veri sayısını temsil etmektedir. Eğitim verimliliğini ve bellek kullanımını etkilemektedir. Büyük parti boyutları, modelin genelleme yeteneğini artırmaktadır ve böylece parametre güncellemelerini daha kararlı hale getirilmektedir. Ayrıca, GPU/TPU gibi donanımın verimli kullanımını sağlayarak eğitim süresini kısaltmaktadır. Daha büyük parti boyutları, iterasyon başına daha fazla veri kullandığından, kayıp fonksiyonu hesaplamalarında ve gradyan tahminlerinde daha az varyans olmaktadır, bu da daha doğru güncellemeler yapılmasını sağlamaktadır. Bunlara ek olarak; daha büyük parti boyutları, daha fazla bellek gerektirmektedir ve bu, özellikle sınırlı belleğe sahip GPU'larda sorun olabilmektedir. Her iterasyonda daha fazla veri işlendiği için, her iterasyonun tamamlanması daha uzun sürebilmektedir ve bu durum eğitim süresini uzatmaktadır. Küçük parti boyutları ise daha az bellek kullanmaktadır, bu da özellikle sınırlı belleğe sahip cihazlarda avantaj sağlamaktadır. Minimum parti boyutun kullanılan donanıma ve veri setine göre karar verilmektedir.

Öğrenme oranı, modelin ağırlıklarının güncellenme hızını belirlemektedir. Öğrenme oranının artması ile model daha hızlı öğrenmektedir, ancak aşırı büyük öğrenme oranları, öğrenme sürecinde kararsızlıklara ve salınımlara yol açabilmektedir. Yüksek öğrenme oranı, modelin optimum çözüme yakınsamasını engelleyebilmektedir ve bu da genellikle modelin genel performansını düşürmektedir. Daha küçük öğrenme oranında ise öğrenme daha yavaş olur, bu da eğitim sürecini uzatabilmektedir. Çok düşük bir öğrenme oranı, modelin eğitim verisine aşırı uyum sağlamasına ve bu nedenle test verisi üzerinde kötü performans göstermesine neden olabilmektedir. Öğrenme oranı parametresine probleme, veri setinin dağılımına ve veri sayısına bakılarak karar verilebilmektedir.

Aktivasyon fonksiyonu, bir nöronun çıktısını belirlemektedir ve modelin doğrusal olmayan ilişkileri öğrenmesini sağlamaktadır. Relu (Rectified Linear Unit) aktivasyon fonksiyonu negatif değerler için 0 değerini, pozitif değerlerin ise kendi değerlerini çıktı olarak vermektedir. Tanh fonksiyonu türevlenebilir ve (-1,1) aralığında değerler alabilen bir fonksiyondur. Elu (Exponential Linear Unit) fonksiyonu ise Relu aktivasyon fonksiyonunun eksikliklerini gidermek için geliştirilmiştir. Negatif değere sahip nöronlara Relu ile 0 değeri atanması bu nöronlarda öğrenme olmayacağı anlamına gelmektedir. Buna alternatif olarak Elu fonksiyonu ile negatif değerlere 0 atamak yerine üstel bir hesaplama yapılmaktadır. Negatif girdilerde Elu fonksiyonunun hesaplama formülü Denklem 1'de görülmektedir. Burada belirtilen  $\alpha$  değeri genellikle 1 olarak kullanılmaktadır. Farklı değerler ile de kullanılabilir.

$$f(x) = \alpha(e^x - 1) \text{ eğer } x \leq 0 \quad (1)$$

Bu parametrelere ek olarak oluşturulan ESA'da 2x2 boyutunda maksimum havuzlama katmanı bulunmaktadır.

**Tablo 1.** Optimize Edilecek Hiper Parametreler ve Değerleri

Hiper parametreler	Seçenekler
Kanal Sayısı	8, 16, 32
Evrışimsel Katman	1x1, 3x3, 5x5, 7x7, 9x9
Minimum Parti Boyutu	32, 64, 128
Öğrenme Oranı	$0 < x \leq 0.5$
Aktivasyon Fonksiyonu	Relu, Tanh, Elu

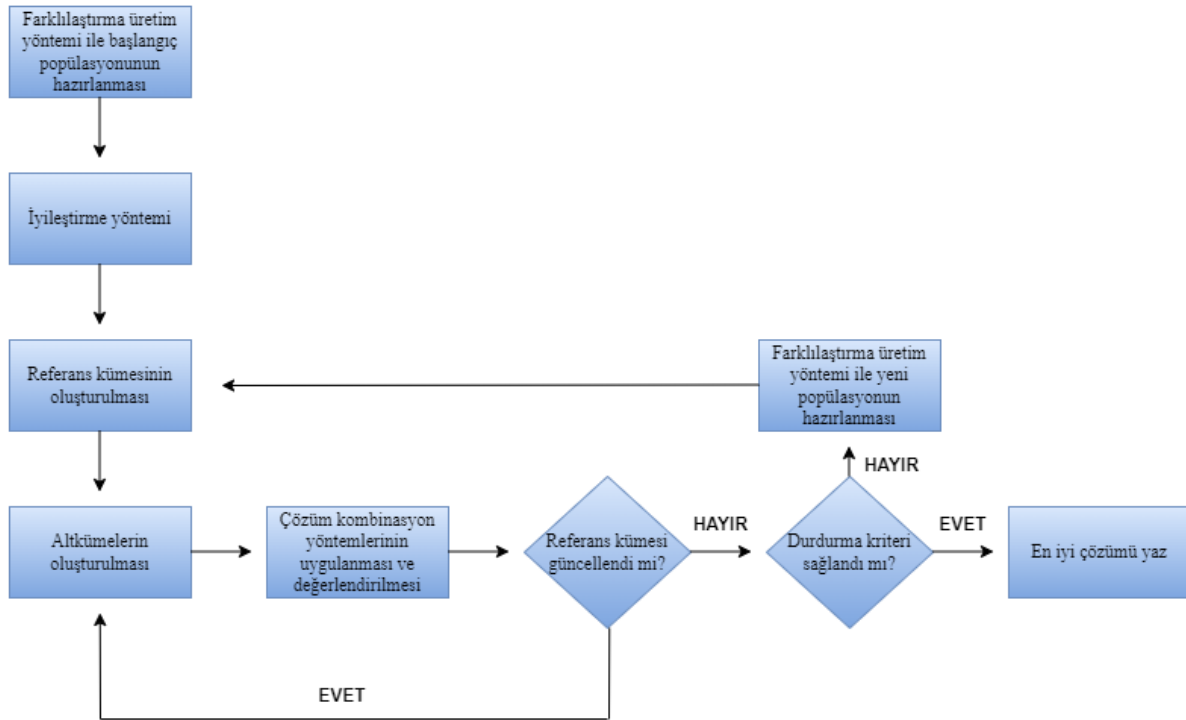
### **Dağınık Arama Algoritması**

DA algoritması ilk olarak 1977'de Glover'ın "Yedek Kısıtlamalar Kullanarak Tamsayı Programlama İçin Meta-sezgisel Yaklaşımlar" başlıklı çalışması ile (Glover, 1977) ortaya atılmıştır. DA, karmaşık optimizasyon problemlerini çözmek için geliştirilmiş bir meta-sezgisel yöntemdir. Geleneksel yöntemlerin aksine DA, çözüm uzayını rastgele aramak yerine, çeşitlilik ve iyileştirme stratejilerini kullanarak daha odaklı bir arama gerçekleştirmektedir. DA algoritmasının diğer evrimsel algoritmalarından farkı ve üstünlüğü, çözüm seçiminde ve yeni çözümler üretme yöntemindedir. DA algoritması, referans seti (RefSet) olarak adlandırılan en iyi çözümlerden oluşan bir set oluşturmaktadır. Bu set, yalnızca uygunluk değeri en yüksek çözümleri değil, aynı zamanda çeşitliliği artırabilecek çözümleri de içermektedir. DA algoritması, referans setindeki çözümleri sistematik olarak seçerek

kombinasyon işlemleriyle yeni çözümler üretmektedir. Yeni çözümler, çözüm havuzuna eklenmektedir ve referans seti güncellenmektedir. Eğer referans setinde değişiklik olmazsa, yeni bir popülasyon oluşturularak bu popülasyondan seçilen bireyler referans setine eklenmektedir. DA algoritması, küresel arama için etkili olup, popülasyonun tamamını işlemeye odaklanmak yerine en başarılı bireyleri ve en farklı bireyleri referans setine ekleyerek çeşitliliği sağlamaya çalışmaktadır (Huang, Wunsch, Levine, & Jo, 2008). Bu yaklaşım, çözümlerin çeşitliliğini koruyarak daha geniş bir çözüm alanında arama yapmayı mümkün kılmaktadır. DA algoritmasının bu özellikleri, GA ve diğer rakip algoritmalara kıyasla daha çeşitli ve potansiyel olarak daha yüksek kaliteli çözümler üretmesini sağlamaktadır. Hazırlanan referans kümesinden alt kümeler oluşturulmaktadır. Alt küme elemanları farklı sayıda belirlenebilmektedir ve bu çalışmada iki elemanlı alt kümeler seçilmiştir. Oluşturulan alt küme sayısının hesaplanma formülü Denklem 2’de görülmektedir. **RFC** terimi, oluşturulan referans kümelerinin sayısını ifade etmektedir.

$$\frac{(RFC * (RFC + 1))}{2} \quad (2)$$

Oluşturulan alt kümelerde yer alan bireylerden çözüm kombinasyon yöntemleri ile çocuk bireyler elde edilmektedir. Elde edilen çocuk bireyler ve referans kümesinde yer alan bireyler arasından RFC adet en iyi çözüm seçilerek yeni referans kümesi hazırlanmaktadır. Eğer seçilen yeni referans kümesi ve eski referans kümesinden farklı ise, tekrar alt küme oluşturulma aşamasına geçilmektedir. Yeni oluşturulan referans kümesi ile eski referans kümesi eşit ise durdurma şartına bağlı olarak referans kümesinin RFC/2 adedi arama uzayında rastgele belirlenen yeni çözümler ile değiştirilmektedir. DA algoritmasının akış şeması Şekil 4’te görülmektedir.



Şekil 4. Dağınık Arama Algoritması Akış Şeması

### ***Dağınık Arama Algoritmasının Hiper Parametre Optimizasyonu için Uygulanması***

Optimize edilecek hiper parametre değerlerine karar verildikten sonra aday çözümler oluşturulmaktadır. Oluşturulan aday çözümlerde rastgele değerler üretilmektedir. Örneğin; kanal sayısı parametresi Tablo 1’de görüldüğü gibi 3 farklı değer alabilmektedir. Bunun için 1 ile 3 arasında rastgele değerler üretilmektedir. Oluşturulan rastgele değerler dönüşüm fonksiyonu ile hiper parametrelerin temsil ettiği değerlere dönüştürülmektedir. Kanal sayısı parametresi için rastgele oluşturulan değer 3 ise dönüşüm fonksiyonunda kanal sayısı parametresine 32 değeri atanmaktadır. Hiper parametrelerden oluşan aday çözümler ve karşılık geldikleri değerlerden bir örnek Tablo 2’de görülmektedir.



**Tablo 2.** Örnek Bir Aday Çözüm ve Karşılık Geldiği Hiper Parametre Değerleri

	Kanal Sayısı	Evrişimsel Katman	Minimum Parti Boyutu	Öğrenme Oranı	Aktivasyon Fonksiyonu
Aday Çözüm	3	1	2	0 – 0.5 arası	3
Karşılık Geldiği Değer	32	1x1	64	0.013	Elu

Aday çözümlerde bulunan değerlerin uygunluklarını hesaplamak için bir ESA yapısı oluşturulmuştur. Oluşturulan ESA'nın eğitimi sonucunda elde edilen doğruluk değeri uygunluk fonksiyonu olarak kullanılmaktadır. Hesaplanan doğruluk değeri uygunluk fonksiyonundan dönen kalite değerini temsil etmektedir.

Bunlara ek olarak çalışmada kullanılacak çaprazlama teknikleri belirlenmiştir. Çalışmada tek nokta çaprazlama, çift nokta çaprazlama ve tekdüze çaprazlama tekniklerinin kullanılmıştır. Her bir çaprazlama tekniğinden 2 çocuk birey olmak üzere toplamda 6 adet çocuk birey oluşturulmuştur. Kullanılan çaprazlama teknikleri şu şekildedir:

**Tek Nokta Çaprazlama:** Rastgele bir çaprazlama noktası belirlenmektedir. Belirlenen çaprazlama noktası P1'den alınmaktadır. Bu nokta dışında kalan kısımlar ise P2'den tamamlanmaktadır. Şekil 5'te örnek bir uygulama görülmektedir.

<b>P1</b>	4	3	2	2	7	2	1	3
<b>P2</b>	3	2	1	2	5	1	2	3

<b>C1</b>	3	2	1	2	7	1	2	3
-----------	---	---	---	---	---	---	---	---

**Şekil 5.** Tek Nokta Çaprazlama Tekniğinden Bir Örnek

**Çift Nokta Çaprazlama:** Rastgele iki tane çaprazlama noktası belirlenmektedir. Belirlenen iki nokta arasındaki değerler P1'den alınmaktadır. Bu iki nokta dışında kalan değerler ise P2'den tamamlanmaktadır. Şekil 6'da örnek bir uygulama görülmektedir.

<b>P1</b>	4	3	2	2	7	2	1	3
<b>P2</b>	3	2	1	2	5	1	2	3

<b>C1</b>	3	3	2	2	7	2	2	3
-----------	---	---	---	---	---	---	---	---

**Şekil 6.** Çift Nokta Çaprazlama Tekniğinden Bir Örnek

**Tekdüze Çaprazlama:** Maskeleye işlemi ile bir çaprazlama yapılmaktadır. 0 ve 1'lerden oluşan bir dizi ile maskeleye yapılmaktadır. 0 değerinde P2'den 1 değerinde P1'den eleman alınmaktadır. Bu şekilde bir çocuk oluşturulmaktadır. Şekil 7'de örnek bir uygulama görülmektedir.

<b>P1</b>	4	3	2	2	7	2	1	3
<b>P2</b>	3	2	1	2	5	1	2	3

<b>Mask</b>	0	1	1	0	0	1	0	1
-------------	---	---	---	---	---	---	---	---

<b>C1</b>	3	3	2	2	5	2	2	3
-----------	---	---	---	---	---	---	---	---

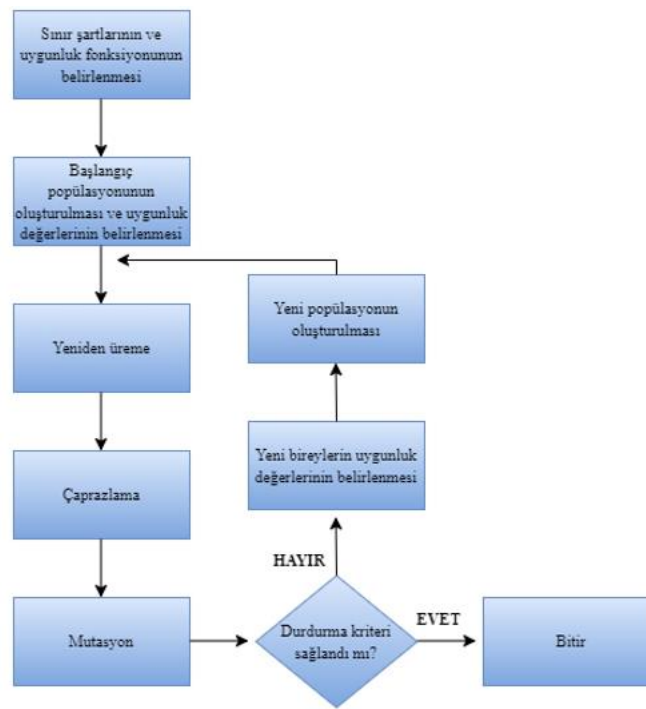
**Şekil 7.** Tekdüze Çaprazlama Tekniğinden Bir Örnek

Belirtilen problem boyutuna göre rastgele tam sayılar ile oluşturulan popülasyonda yinelenen değerler çıkarılarak eşsiz bireyler ile çalışılmıştır. Rastgele olarak oluşturulan bireyler uygunluk fonksiyonuna gönderilerek ESA'dan bir doğruluk değeri elde edilmiştir. Elde edilen en iyi ve en farklı sonuçlar kaydedilmiştir. Farklı değerler ile sonuçlar elde etmek için belirtilen çaprazlama teknikleri kullanılmıştır. Çaprazlama tekniklerinin farklı çalışma prensipleri ile popülasyonda farklı bireyler oluşturulmuştur. Çalışmanın isterleri tamamlandıktan sonra test edilme aşamasına geçilmiştir. Test aşamasında aynı hiper parametre değerlerine karşılık gelen farklı doğruluk değerleri elde edilmiştir. Bu tutarsızlık sorununu çözmek için ESA'nın ağırlık değerlerinin sabitlenmesi düşünülmüştür. ESA ağırlıkları

rng(123) komutu ile sabitlenmiştir. Bu şekilde aynı parametreler her eğitildiğinde aynı doğruluk değerini vermektedir. Fakat bu komutu ekledikten sonra popülasyon oluşumunda rassallık bozulduğu için rng('shuffle') komutu projeye eklenerek rassallık korunmuştur.

### Genetik Algoritma

GA ilk olarak 1976 yılında John Holland tarafından ortaya atılmıştır. Holland, GA'nın teorik temellerini "Adaptation in Natural and Artificial Systems" adlı kitabında tanıtmıştır (Sampson, 1976). Bu kitap, evrimsel süreçlerin bilgisayar sistemlerinde nasıl uygulanabileceğini açıklayan ilk kapsamlı çalışmadır. John Holland, Michigan Üniversitesi'nde biyolojik evrim sürecinden esinlenerek doğal seleksiyon ilkesine dayalı bir hesaplama yöntemi geliştirmiştir. GA, özellikle karmaşık problemlerin çözümünde kullanılan, doğal seleksiyon, mutasyon ve çaprazlama gibi biyolojik süreçleri simüle eden bir optimizasyon yöntemidir. GA'nın başarısı, popülasyon büyüklüğü, jenerasyon sayısı, mutasyon ve çaprazlama oranları gibi hiper parametrelerin dikkatlice ayarlanmasına bağlıdır. Hiper parametre optimizasyonu, algoritmanın çözüm kalitesini ve işlem süresini doğrudan etkilemektedir. Şekil 8'de GA'nın akış şeması görülmektedir.



Şekil 8. GA Akış Şeması

Çalışmada, DA algoritması ve GA arasında kapsamlı bir karşılaştırma gerçekleştirilmiştir. Bu karşılaştırmanın sağlıklı ve adil olabilmesi için, her iki algoritma aynı parametreler, veri seti ve çaprazlama teknikleri kullanılarak çalıştırılmıştır. Böylece, her iki yöntemin performansını etkileyen faktörlerin eşitlenmesi sağlanmış ve daha anlamlı sonuçlar elde edilmiştir.

Veri seti, her iki algoritmanın genel performansını değerlendirmek için dikkatlice seçilmiş olup, hem GA hem de DA algoritması için optimize edilmiş sonuçların karşılaştırılmasına olanak tanımaktadır. Kullanılan çaprazlama teknikleri, algoritmaların çözüm uzayını keşfetme yeteneklerini artırmak amacıyla belirlenmiş ve her iki yöntemin farklı stratejilerini test etme fırsatı sunmuştur.

### BULGULAR ve DENEYSEL SONUÇLAR

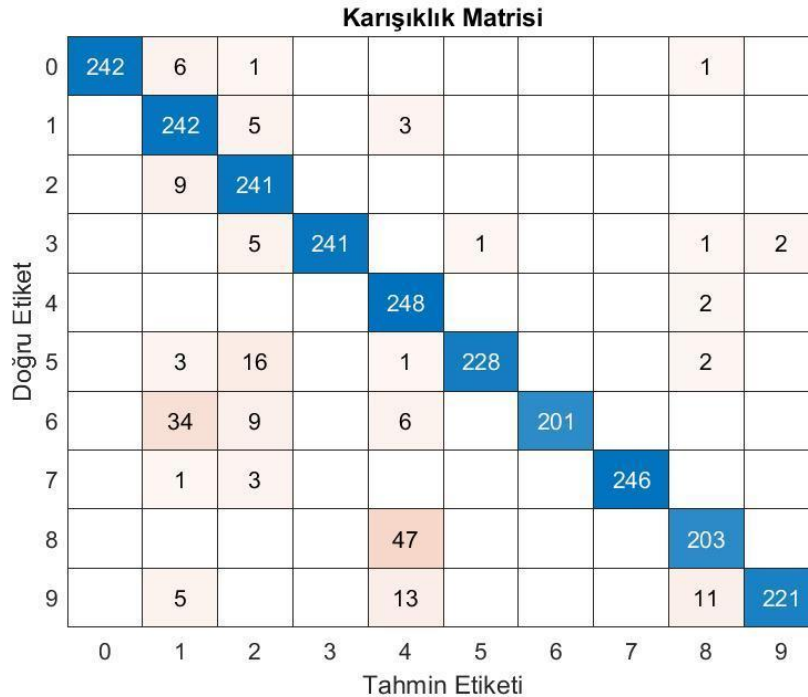
Projenin geliştirilme aşamasında GA ve DA algoritması için popülasyon boyutu (NP) 20, maksimum FEs değeri 1000, çalışma zamanı değeri 5 ve ESA devir değeri 10 olarak belirlenmiştir. Geng vd. (2016) çalışmasında, optimizasyon algoritmalarında popülasyon boyutunun probleme göre belirlendiği ve genellikle 20 - 50 arasında bir değer verildiği belirtilmiştir. Bu doğrultuda, optimize edilecek problem ve kullanılacak donanımın algoritmaların çalışma süresine etkisi göz önünde bulundurularak NP değerinin 20 olarak belirlenmesine karar verilmiştir. FEs değeri, algoritmanın çalışması sırasında değerlendirilen toplam çözüm sayısını ifade etmektedir. Her bir çözüm,

algoritmanın belirli bir problem için uygunluk veya maliyet fonksiyonunu değerlendirerek ne kadar iyi bir çözüm olduğunu belirlemektedir. Maksimum FEs değeri, algoritmanın ne kadar derinlemesine arama yapacağını ve dolayısıyla çözüm uzayını ne kadar kapsamlı bir şekilde keşfedeceğini belirlemektedir. Algoritmaların çalışma süresi ve donanım özellikleri dikkate alınarak, dengeli bir performans ve verimlilik sağlamak amacıyla maksimum FEs değeri 1000 olarak belirlenmiştir. Algoritmaların çalıştırılması ile en başarılı sonuç DA algoritmasından elde edilmiştir. DA algoritmasından elde edilen en yüksek uygunluk değeri yani ESA'dan elde edilen doğruluk değeri % 93.24'tür. GA'dan elde edilen en yüksek doğruluk değeri ise %88.76'dır. DA algoritması ve GA için en yüksek doğruluk değerlerini veren hiper parametre değerleri Tablo 3'te görülmektedir.

**Tablo 3.** En Yüksek Doğruluk Değerinin Elde Edildiği Hiper Parametre Değerleri

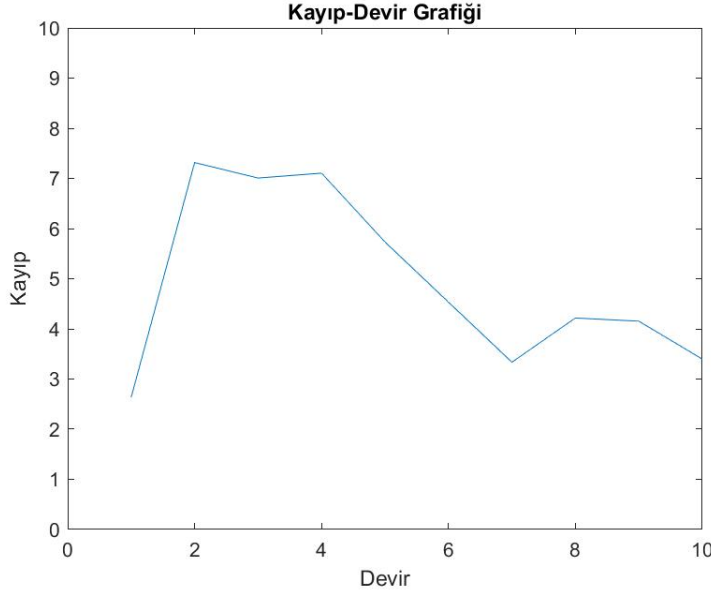
	Kanal Sayısı	Evrimsel Katman	Minimum Parti Boyutu	Öğrenme Oranı	Aktivasyon Fonksiyonu
(DA) Aday Çözüm	2	3	2	0 - 0.5 arası	1
(DA) Karşılık Geldiği Değer	16	5x5	64	0.0052	reluLayer
(GA) Aday Çözüm	3	2	2	0 - 0.5 arası	1
(GA) Karşılık Geldiği Değer	32	3x3	64	0.0185	reluLayer

En başarılı sonucun elde edildiği DA algoritması ile ESA modelinin başarısını değerlendirmek için, MNIST veri setindeki test verileri kullanılarak bir karışıklık matrisi hesaplanmıştır. Elde edilen karışıklık matrisi Şekil 9'da görülmektedir. Karışıklık matrisinde köşegenlerde bulunan değerler doğru tahminlerin sayısını ifade etmektedir. Karışıklık matrisi, sınıflandırma algoritmalarının performansını değerlendirmek için kullanılan bir metriktir. Bu matris, tahmin edilen sınıfların gerçek sınıflarla karşılaştırılmasını sağlamaktadır. Karışıklık matrisi kullanılarak, doğruluk, hassasiyet (precision), duyarlılık (recall) ve F1-skor değerleri hesaplanabilmektedir. Hassasiyet, pozitif olarak tahmin edilen örneklerin gerçekten pozitif olma oranıdır. Duyarlılık, gerçekten pozitif olan örneklerin ne kadarının pozitif olarak tahmin edildiğini göstermektedir. F1-skor, hassasiyet ve duyarlılık arasındaki dengeyi gösteren bir metriktir özellikle dengesiz olarak dağılmış veri setlerinde tercih edilmektedir. Karışıklık matrisine bakıldığında en başarılı sınıflandırma işleminin en fazla doğru tespit ile "4" sınıfına ait olduğu görülmektedir.



**Şekil 9.** Karışıklık Matrisi

Kayıp değeri, modelin tahminlerinin gerçek değerlerden ne kadar farklı olduğunu ölçen bir metriktir. 10 devir ile gerçekleştirilen eğitim sürecinde kayıp değerinin her devirdeki değişimi Şekil 10'daki grafikte görülmektedir.



Şekil 10. Kayıp Değerinin Değişim Grafiği

## SONUÇ

Bilgisayarlı görü çalışmalarında hiper parametrelerin aldığı değerler elde edilen sonucu doğrudan etkilemektedir. Buradan yola çıkarak çalışmada çoklu sınıflandırma problemi için kullanılacak bir ESA tasarlanmıştır. ESA'dan verimli bir sonuç elde edebilmek için belirlenen bazı hiper parametrelerin optimizasyonu gerçekleştirilmiştir. Optimizasyon için literatürde yaygın olan algoritmalarından farklı olarak DA algoritması kullanılarak ESA ağının hiper parametre optimizasyonu gerçekleştirilmiştir. Çalışmanın özgünlüğü, DA algoritmasının ESA gibi derin öğrenme modellerinin hiper parametre optimizasyonunda kullanılmasıdır. Literatürde, GA, PSO ve yapay arı kolonisi gibi çeşitli evrimsel algoritmalar yaygın olarak kullanılırken, DA algoritmasının bu alandaki uygulaması sınırlı kalmaktadır. Buradan yola çıkılarak DA algoritması aynı parametreler, donanım ve veri seti üzerinden GA ile karşılaştırılmıştır. Yapılan testler sonucunda DA algoritmasının GA'dan daha başarılı sonuçlar verdiği gözlemlenmiştir. Bu çalışma, DA algoritmasının hiper parametre optimizasyonunda etkili bir şekilde kullanılabileceğini göstererek literatüre önemli bir katkı sağlamaktadır. Ayrıca, çalışmada üç farklı çaprazlama tekniği kullanılmıştır. Bu yaklaşım, özellikle geniş çözüm alanlarında daha çeşitli ve yüksek kaliteli çözümler üretme potansiyeli sunarak, derin öğrenme modellerinin performansını artırma konusunda yeni bir perspektif sunmaktadır. Gerçekleştirilen optimizasyon sonucunda en iyi sonucu veren hiper parametreler belirlenerek ESA eğitiminde kullanılmıştır.

Gelecekteki çalışmalarda daha fazla çaprazlama tekniği ile farklı aday çözümler oluşturulması, kullanılan donanıma bağlı olarak ayrık ve sürekli değerlerin sayısı artırılması ve daha fazla hiper parametrenin optimize edilmesi önerilmektedir.

## KAYNAKLAR

- Ait Amou, M., Xia, K., Kamhi, S., & Mouhafid, M. (2022, March). A novel MRI diagnosis method for brain tumor classification based on CNN and Bayesian Optimization. In Healthcare (Vol. 10, No. 3, p. 494). MDPI.
- Andonie, R., & Florea, A.-C. (2020). Weighted random search for CNN hyperparameter optimization. arXiv preprint arXiv:2003.13300.
- Anguita, D., Ghio, A., Oneto, L., Parra, X., & Reyes-Ortiz, J. L. (2013). A public domain dataset for human activity recognition using smartphones. Paper presented at the Esann.

- Aslan, M. F., Sabanci, K., Durdu, A., & Unlarsen, M. F. (2022). COVID-19 diagnosis using state-of-the-art CNN architecture features and Bayesian Optimization. *Computers in biology and medicine*, 142, 105244.
- Atteia, G., Abdel Samee, N., El-Kenawy, E. S. M., & Ibrahim, A. (2022). CNN-hyperparameter optimization for diabetic maculopathy diagnosis in optical coherence tomography and fundus retinography. *Mathematics*, 10(18), 3274.
- Bochinski, E., Senst, T., & Sikora, T. (2017). Hyper-parameter optimization for convolutional neural network committees based on evolutionary algorithms. Paper presented at the 2017 IEEE international conference on image processing (ICIP).
- Bozkurt, F., & Yağanoğlu, M. (2021). Derin evrişimli sinir ağları kullanarak akciğer X-Ray görüntülerinden COVID-19 tespiti. *Veri Bilimi*, 4(2), 1-8.
- Fujino, S., Mori, N., & Matsumoto, K. (2017). Deep convolutional networks for human sketches by means of the evolutionary deep learning. Paper presented at the 2017 joint 17th world congress of international fuzzy systems association and 9th international conference on soft computing and intelligent systems (IFSA-SCIS).
- Geng, J. C., Cui, Z., & Gu, X. S. (2016). Scatter search based particle swarm optimization algorithm for earliness/tardiness flowshop scheduling with uncertainty. *International Journal of Automation and Computing*, 13(3), 285-295.
- Glover, F. (1977). Heuristics for integer programming using surrogate constraints. *Decision sciences*, 8(1), 156-166.
- Gülcü, A., & Kuş, Z. (2019). A survey of hyper-parameter optimization methods in convolutional neural networks. *Gazi Üniversitesi Fen Bilimleri Dergisi*, 7(2), 503-522.
- Hao, W., Yizhou, W., Yaşin, L., & Zhili, S. (2020, December). The role of activation function in CNN. In 2020 2nd International Conference on Information Technology and Computer Application (ITCA) (pp. 429-432). IEEE.
- Huang, D. S., Wunsch, D. C., Levine, D. S., & Jo, K. H. (Eds.). (2008). *Advanced Intelligent Computing Theories and Applications. With Aspects of Artificial Intelligence: Fourth International Conference on Intelligent Computing, ICIC 2008 Shanghai, China, September 15-18, 2008, Proceedings (Vol. 5227)*. Springer Science & Business Media.
- İnik, Ö. (2023). CNN hyper-parameter optimization for environmental sound classification. *Applied Acoustics*, 202, 109168.
- Kıymaç, M. E. (2022). Hyper-parameter optimization of deep neural networks with metaheuristic algorithms. *Yüksek Lisans Tezi, Alparslan Türkeş Bilim ve Teknoloji Üniversitesi*.
- Lorenzo, P. R., Nalepa, J., Kawulok, M., Ramos, L. S., & Pastor, J. R. (2017). Particle swarm optimization for hyper-parameter selection in deep neural networks. Paper presented at the Proceedings of the genetic and evolutionary computation conference.
- MNIST Dataset. (2010). <http://yann.lecun.com/exdb/mnist>.
- Mooney, P. T. Retinal OCT images (Optical coherence tomography). Kaggle. (2018). <https://www.kaggle.com/datasets/paultimothymooney/kermany2018>
- Özbay, E., & Özbay, F. A. (2023). Parçacık Sürüşü Optimizasyon Algoritması ile Optimize Edilmiş Evrişimsel Sinir Ağı Kullanılarak Dermoskopik Görüntülerden Cilt Kanserinin Sınıflandırılması. *Fırat Üniversitesi Mühendislik Bilimleri Dergisi*, 35(1), 261-273.
- Piczak, K. J. (2015). ESC: Dataset for environmental sound classification. Paper presented at the Proceedings of the 23rd ACM international conference on Multimedia.
- Porwal, P.; Pachade, S.; Kamble, R.; Kokare, M.; Deshmukh, G.; Sahasrabudhe, V.; Meriaudeau, F. Indian Diabetic Retinopathy Image Dataset (IDRiD): A Database for Diabetic Retinopathy Screening Research. *Data* 2018, 3, 25.
- Raziani, S., & Azimbagirad, M. (2022). Deep CNN hyperparameter optimization algorithms for sensor-based human activity recognition. *Neuroscience Informatics*, 2(3), 100078.
- Sakib, S., Ahmed, N., Kabir, A. J., & Ahmed, H. (2019). An overview of convolutional neural network: Its architecture and applications.



Salamon, J., Jacoby, C., & Bello, J. P. (2014). A dataset and taxonomy for urban sound research. Paper presented at the Proceedings of the 22nd ACM international conference on Multimedia.

Sampson, J. R. (1976). *Adaptation in natural and artificial systems* (John H. Holland).

Sethi, M., Ahuja, S., Rani, S., Bawa, P., & Zaguia, A. (2021). [Retracted] Classification of Alzheimer's Disease Using Gaussian-Based Bayesian Parameter Optimization for Deep Convolutional LSTM Network. *Computational and Mathematical Methods in Medicine*, 2021(1), 4186666.

Shankar, K., Zhang, Y., Liu, Y., Wu, L., & Chen, C. H. (2020). Hyperparameter tuning deep learning for diabetic retinopathy fundus image classification. *IEEE access*, 8, 118164-118173.

Tanyıldızı, E., & Demirtaş, F. (2019). Hiper Parametre Optimizasyonu Hyper Parameter Optimization. Paper presented at the 2019 1st International Informatics and Software Engineering Conference (UBMYK).

Yurdakul, M. (2022). Meta-sezgisel algoritmalar ile konvolüsyonel sinir ağı mimarisinin hiper parametrelerinin optimizasyonu. Yüksek Lisans Tezi, Selçuk Üniversitesi.

ZainEldin, H., Gamel, S. A., El-Kenawy, E. S. M., Alharbi, A. H., Khafaga, D. S., Ibrahim, A., & Talaat, F. M. (2022). Brain tumor detection and classification using deep learning and sine-cosine fitness grey wolf optimization. *Bioengineering*, 10(1), 18.

Zhang, B., Rajan, R., Pineda, L., Lambert, N., Biedenkapp, A., Chua, K., ... & Calandra, R. (2021, March). On the importance of hyperparameter optimization for model-based reinforcement learning. In *International Conference on Artificial Intelligence and Statistics* (pp. 4015-4023). PMLR.

Zhang, M., Li, H., Pan, S., Lyu, J., Ling, S., & Su, S. (2021). Convolutional neural networks-based lung nodule classification: A surrogate-assisted evolutionary algorithm for hyperparameter optimization. *IEEE Transactions on Evolutionary Computation*, 25(5), 869-882.

Xiao, X., Yan, M., Basodi, S., Ji, C., & Pan, Y. (2020). Efficient hyperparameter optimization in deep learning using a variable length genetic algorithm. *arXiv preprint arXiv:2006.12703*.