

PAPER DETAILS

TITLE: Beton atlakların derin ğrenme tabanlı semantik segmentasyonunda kodlayıcı
değişkenlerinin karşılaştırmalı analizi

AUTHORS: Hasan Polat,Serhat Alpergin,Mehmet Sıra Özerdem

PAGES: 581-593

ORIGINAL PDF URL: <https://dergipark.org.tr/tr/download/article-file/3850850>



Araştırma Makalesi / Research Article

Beton çatlakların derin öğrenme tabanlı semantik segmentasyonunda kodlayıcı değişkenlerinin karşılaştırmalı analizi

Comparative analysis of encoder variants in deep learning-based semantic segmentation of concrete cracks

Hasan Polat^{1*}, Serhat Alpergin², Mehmet Siraç Özerdem³

¹ Bingöl Üniversitesi, Elektrik ve Enerji Bölümü, hasanpolat@bingol.edu.tr
ORCID: <https://orcid.org/0000-0001-5535-4832>

² Dicle Üniversitesi, Elektrik Elektronik Mühendisliği Bölümü, serhatalpergin@gmail.com
ORCID: <https://orcid.org/0009-0009-7780-772X>

³ Dicle Üniversitesi, Elektrik Elektronik Mühendisliği Bölümü, sozerdem@dicle.edu.tr
ORCID: <https://orcid.org/0000-0002-9368-8902>

MAKALE BİLGİLERİ

Makale Geçmişi:

Geliş 5 Nisan 2024
Revizyon 1 Ağustos 2024
Kabul 15 Ağustos 2024
Online 30 Eylül 2024

Anahtar Kelimeleri:

Semantik Segmentasyon, Derin Öğrenme, Görüntü İşleme, Çatlak Tespiti, Çatlak Segmentasyonu, Yapı Denetimi

ÖZ

Depremler, seller ve yangınlar gibi doğal afetler akabinde yapılarda ve kentsel altyapıda ciddi hasarlar meydana gelmektedir. Çatlaklar, beton yapılarda meydana gelen hasarların veya bozulmaların en yaygın belirtileri olarak kabul edilmektedir. Dolayısıyla, çatlak kusurlarının erken ve doğru bir şekilde tespit edilmesi, bu tür yapıların güvenliklerinin sağlanması ve hizmet süreleri açısından önem arz etmektedir. Son yıllarda bilgisayarlı görü uygulamalarında önemli bir atılım sergileyen derin öğrenme mimarileri, beton çatlaklarının otomatik olarak tespit ve segmente edilmesinde yaygın olarak kullanılmaya başlanmıştır. Özellikle, genel olarak bir kodlayıcı ve bir kod çözücü bloktan oluşan derin öğrenme tabanlı segmentasyon mimarileri çatlakları uzamsal sınırları ile tespit ederek, kapsamlı yapı sağlığı analizlerini mümkün kılmaktadır. Ancak, evrimsel filtrede küçük alıcı alan, pooling işleminin neden olduğu bilgi kayıpları ve yetersiz yerel özellik işlenmesi gibi kodlayıcı blok sınırlandırmaları segmentasyon performansını sekteye uğratmaktadır. Bu çalışmada, beton yüzeylerindeki çatlakların segmentasyonu için önerilen DeepLabV3+ mimarisinde kodlayıcı blok için farklı omurga mimarilerinin (ResNet-18, ResNet-50, MobileNetV2, Xception ve Inception-ResNet) etkinlikleri analiz edilmiştir. Farklı omurga mimariler ile sağlanan alçak ve yüksek seviyeli özelliklerin etkinliklerinin test edilmesi için erişime açık DeepCrack ve CrackForest veri setleri kullanılmıştır. Bulgular her iki veri seti için de MobileNetV2 mimarisinin eğitilebilir parametre ve segmentasyon performansı açısından en başarılı ağ olduğunu göstermiştir. MobileNetV2 kodlayıcı tabanlı segmentasyon çerçevesi, yaklaşık 6.7 milyon eğitilebilir ağırlık kullanılarak her iki veri seti için sırasıyla 0.81 ve 0.70 Dice benzerlik katsayısı (DSC) başarımları elde etmiştir. Deneysel bulgular, kodlayıcı bloğunun derinliğini ve karmaşıklığını artırmanın segmentasyon doğruluğunu artırmadığını, oysa daha sıkı özellik çıkarıcıların beton çatlaklarını tespit etmede daha etkili olduğunu göstermektedir.

ARTICLE INFO

Article history:

Received 5 April 2024
Received in revised form 1 August 2024
Accepted 15 August 2024
Available online 30 September 2024

Keywords:

Semantic Segmentation, Deep Learning, Image Processing, Crack Detection, Crack Segmentation, Building inspection

Doi: 10.24012/dumf.1465724

* Sorumlu Yazar

ABSTRACT

Following natural disasters such as earthquakes, floods, and fires, significant damages manifest in both buildings and urban infrastructure. Cracks are widely recognized as the predominant indicators of damage or deterioration in concrete structures. Hence, the early and accurate detection of crack defects is crucial to ensure structural safety and longer service life. Deep learning architectures, which have made a significant breakthrough in computer vision applications in recent years, have begun to be widely used in the automatic detection and segmentation of concrete cracks. In particular, deep learning-based segmentation architectures, typically comprising an encoder and a decoder part, play a crucial role in conducting thorough structural health analyses by precisely detecting cracks along with their spatial boundaries. However, encoder block limitations such as the small receptive field of convolution kernels, information losses caused by the pooling operation, and insufficient local feature processing can hinder segmentation performance. This study examines the efficacy of various backbone architectures (ResNet-18, ResNet-50, MobileNetV2, Xception, and Inception-ResNet) as employed in the encoder block within the DeepLabV3+ framework, proposed for the segmentation of cracks on concrete surfaces. The effectiveness of low-level and high-level features provided by different backbone architectures in the encoder part was evaluated on open-access DeepCrack and CrackForest datasets. The results revealed that the MobileNetV2 architecture was the most successful network in terms of learnable parameters and segmentation performance for both data sets. The MobileNetV2 encoder-based segmentation framework achieved 0.81 and 0.70 Dice similarity coefficient (DSC) for both datasets, respectively, using approximately 6.7 million learnable weights. Experimental results indicate that increasing the depth and complexity of the encoder block does not enhance segmentation accuracy, whereas shallower feature extractors are more effective for detecting concrete cracks.

Giriş

Beton, düşük maliyeti, iyi dayanıklılığı ve şekillendirme kolaylığı nedeniyle bina, otoyol ve köprü gibi altyapı ve üstyapılarda en yaygın olarak kullanılan inşaat malzemesidir [1], [2]. Mevcut yapıların uzun yıllar sürekliliğini koruması için beton kaynaklı deformasyonlara ilişkin belirtiler önem arz etmektedir. Bu kapsamda, çatlaklar beton yapılarda meydana gelen hasarların veya bozulmaların en yaygın belirtileri olarak kabul edilmektedir [2], [3]. Ayrıca, doğal afetler akabinde çökme riskinin ilk sinyali olarak değerlendirilmektedirler [4]. Çatlaklar doğrudan yapıların güvenliğini, uygulanabilirliğini ve dayanıklılığını etkilemektedir. Dolayısıyla, yapılara ilişkin beton yüzeylerde meydana gelen çatlakların yapı denetimi ile hızlı ve doğru bir şekilde tespit edilmesi gerekmektedir.

Yapı denetiminde, uzmana bağlı görsel inceleme sıklıkla uygulanan bir hasar tespit yöntemidir [5]. Bu analiz yönteminde çatlaklar genel olarak çıplak gözle, el tipi detektörlerle veya iskele tipi platform üzerinde çeşitli testler ile yapılır. Ancak, çatlakları tespit etme potansiyeline rağmen, manuel inceleme yoğun emek gerektirir, tehlikelidir, öznel ve zaman alıcıdır [2]. Bu nedenle bilgisayar destekli otomatik hasar tespit sistemlerinin ivedilikle geliştirilmesi önem arz etmektedir.

Yapı sağlığı denetçilerine çatlak tespitinde yardımcı olmak için bilgisayarlı görsel tabanlı sınıflandırma veya segmentasyon amaçlı çok çeşitli otonom destek modelleri geliştirilmiştir. Sınıflandırma çalışmalarında çatlak bulunan görüntüler doğrudan hasarlı olarak işaretlenirken, segmentasyon çalışmalarında ise çatlakların uzamsal sınırları ile birlikte tespit edilmesi mümkün kılınmıştır. Segmentasyon çalışmaları ile çatlaklara ilişkin daha kapsamlı bilgilerin çıkarılması, bu alana olan yönelimi artırmıştır. Günümüze kadar süregelen segmentasyon çalışmalarının ise genel olarak konvansiyonel ve derin öğrenme tabanlı olmak üzere iki sınıf altında toparlanabildiği rapor edilmiştir [6].

Konvansiyonel yaklaşımlarda, çatlaklara ilişkin görüntülerden özelliklerin manuel olarak çıkarılması ve çıkarılan özelliklerin makine öğrenmesi için uygun girdi vektörlerine dönüştürülmesi genel çerçeve olarak yürütülmüştür. Çatlak segmentasyonunda özellikle geometrik ve fotometrik özelliklerden istifade edilmiştir [7]. Ayrıca, kenar algılama, dalgacık dönüşümü ve gri ölçekli histogram gibi temel görüntü işleme teknikleri, çatlakların tespit veya segmentasyonunda sıklıkla başvurulan yaklaşımlar olmuştur [5]. Bu kapsamda, Talab ve ark. [8] beton yüzeylerine ilişkin görüntülerden çatlakların otomatik olarak sınıflandırmasında Sobel filtreleme yöntemi önermişlerdir. Chun ve ark. [9] ise özellik vektörlerinin oluşturulmasında ışık gradyan artırma makinesine dayalı somut piksel değerlerine ve diğer geometrik özelliklere bağlı çatlak algılama modeli önermişlerdir. Zou ve ark. [10] çatlak tespitinde üç aşamadan oluşan bir konvansiyonel yaklaşım uygulamışlardır. Bunlar sırasıyla jeodezik giderme, tensör

oylaması ve kenar algılamadan oluşmuştur. Literatürde önerilen konvansiyonel segmentasyon çerçeveleri belirli bir veri tabanı veya sahne için özel tasarlandığından, önerilen modellerin genelleyici olma ve gerçek dünya problemlerinde uygulanabilirlik kabiliyetlerini sınırlı kılmaktadır [11]. Ayrıca, düşük kontrast, olumsuz aydınlatma koşulları, kirlilik, topaklanma, morfolojik farklılıklar ve teknolojik ilerlemeler ile birlikte büyük veri setleri konvansiyonel yöntemlerin başarımlarını olumsuz etkilemektedir [2], [12]. Dolayısıyla, gerçek dünya problemleri için ilave ön işleme yöntemlerine ve manuel özellik çıkarımına gereksinim duymayan daha pratik çözümlere ihtiyaç duyulmuştur.

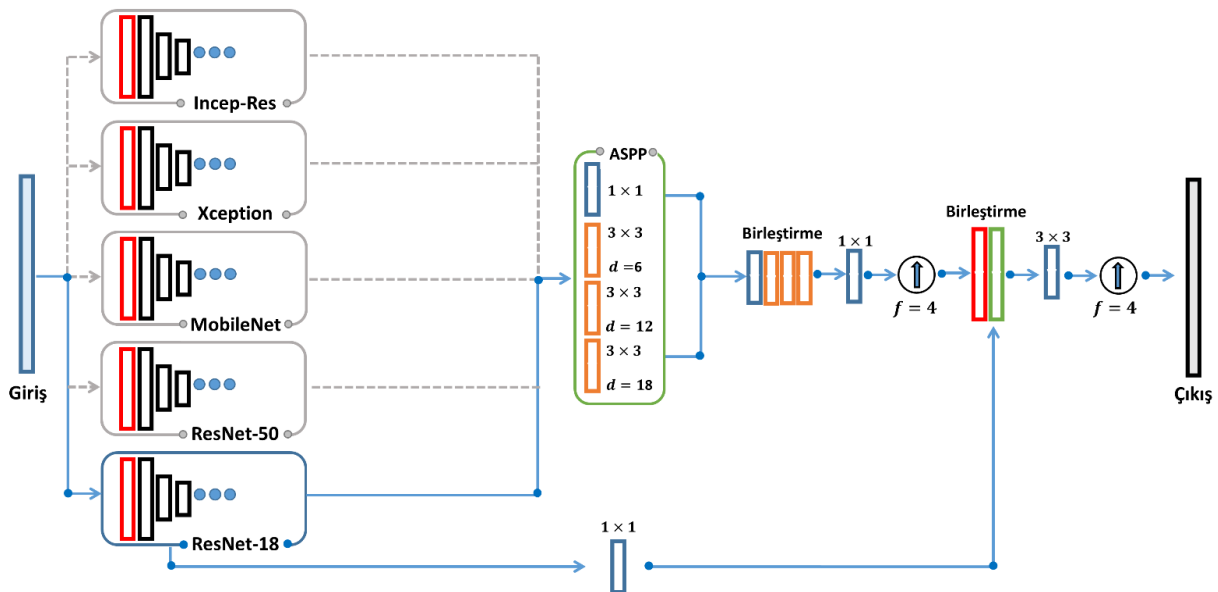
Derin öğrenme son yıllarda birçok alanda dikkat çekici bir atılım sergilemiştir. Ham görüntülerden ayırt edici özellikleri otomatik olarak çıkarabilmeleri ve büyük veri kümelerindeki üstün performansları, derin mimarileri konvansiyonel yöntemlere göre avantajlı kılmıştır [13]. Çatlaklara ilişkin bilgisayarlı görsel uygulamalarında da hızla gelişim göstererek, yaygın olarak kullanılmaya başlanmıştır [2], [14], [15]. Literatürde, derin öğrenme tabanlı önerilen modeller ile yapı görüntülerinden çatlaklara ilişkin sınıflandırma [16], hasarlı nesne tespiti [17] ve semantik segmentasyon [18], [19] uygulamaları başarı ile yürütülmüştür. Ancak, çatlak desenlerinin genellikle ince, belirsiz koyu şeritlerden oluşması veya sürekli olarak değişen açı ve yönlerle sahip çizgilerden oluşması, çatlakların kutu sınırlar veya genel bir örüntü olarak değerlendirilmesini verimsiz kılmıştır. Bu sınırlandırmaların üstesinden gelmek için çatlakların piksel düzeyinde analiz edilmesine ve semantik segmentasyonuna olan yönelimi artırmıştır.

Derin öğrenmenin görüntü işleme üzerindeki kabiliyetlerinden bu alanda istifade edilmesi için çok sayıda segmentasyon çerçevesi geliştirilmiştir. Önerilen segmentasyon çerçevelerinin genellikle U-Net [20], SegNet [21], FCN [22] veya DeepLabV3+ [23] segmentasyon mimarilerinin doğrudan veya mevcut mimariler üzerinde çeşitli revizyonlarının uygulanarak geliştirildiği gözlemlenmiştir. Bu kapsamda, Zhang ve ark. [24] U-Net mimarisinin beton yüzeylerindeki çatlaklara ilişkin kenar bilgisini çıkarmadaki eksikliğini üstesinden gelmek için her kodlayıcı katmanına bir dikkat modeli (attention module) ekleyerek yeni bir model geliştirmişlerdir. ALP-UNet olarak adlandırdıkları model ile kod çözücü modül de geliştirilmiş ve segmentasyon performansında dikkat çekici bir artış gözlemlenmiştir. Kong ve ark. [7] kodlayıcı ve kodçözücü bloktan oluşan ve farklı kombinasyonlarla karşılaştırdıkları bir semantik segmentasyon modeli önermişlerdir. Önerdikleri model ile çatlak segmentasyonunda transfer öğrenimin performansı doğrudan artırdığını gözlemişlerdir. Ayrıca koldayıcı blokta kullanılan farklı omurga mimarilerine ilişkin segmentasyon performanslarını da karşılaştırmışlardır. Omurga mimarilerin değişimiyle optimum görüntü ölçeğinde yaklaşık olarak %4 oranında bir değişim olduğunu tespit etmişlerdir. Lee ve ark. [25] beton yüzey görüntülerinden çatlakların segmentasyonunda mevcut derin modellerin

performanslarını artırmaya çalışmışlardır. Bu kapsamda, U-Net, DeepLabV3+, DANet ve FCN segmentasyon mimarilerini kullanmışlardır. Mevcut mimariler ile segmentasyon başarımları Jaccard indeksi türünden 0.4 ila 0.6 arasında bir değişiklik gösterirken, yığın öğrenme modelinin entegre edilmesi ile performansta ciddi artışların olduğunu gözlemlemişlerdir. Al-Huda ve ark. [26] çatlakların semantik segmentasyonunda karşılaşılan gürültü tabanlı zorlukların üstesinden gelmek için derin öğrenme tabanlı hibrit bir model önermişlerdir. Önerdikleri model ile sınıflandırıcı ve segmentasyon ağı arasında bilgi aktarımını gerçekleştirerek, segmentasyon performansını artırmaya çalışmışlardır. Yang ve ark. [6] derin öğrenme tabanlı segmentasyon mimarilerinde yetersiz yerel özellik işleme ve havuzlama katmanı nedeniyle bilgi kaybı gibi dezavantajların üstesinden gelmek için yeni bir segmentasyon modeli önermişlerdir. U-Net mimarisine ilerici ve hiyerarşik bağlam füzyonu olarak adlandırılan yeni modüller ekleyerek, çatlakların daha yüksek hassasiyetle segmentasyonunu amaçlamışlardır. Sonuç olarak önerdikleri modelin daha üstün bir performans sergilediğini nicel bulgular ile kanıtlamışlardır. Soni ve ark. [27] FCN mimari tabanlı bir çatlak tespit modeli önermişlerdir. Kodlayıcı yapı için ise InceptionV3, ResNet ve VGG16 evrimsel modellerini kullanarak, çıkarılan özelliklerin model performansına etkilerini incelemişlerdir. Bulgular VGG16'nın daha yüksek bir performansla semantik özellikleri çıkardığını göstermiştir. Liu ve ark. [28] beton çatlaklarının segmentasyonunda iki boyutlu uzamsal ve üç boyutlu geometrik bilgilerin bileşime dayanan bir model önermişlerdir. Çok boyutlu füzyon yaklaşımları ile son teknoloji modellerin üzerinde bir bir performans elde etmişlerdir. Lang ve ark. [29] yapay zeka tabanlı bir dikkat mekanizması ile etkinliği geliştirilen bir derin segmentasyon mimarisi önermişlerdir. Önerdikleri model ile çatlak dışındaki arak plan girişimlerinin performans üzerindeki olumsuz etkiye önlenmeye çalışılmıştır.

Literatürden de görüldüğü üzere, çatlak segmentasyonunda derin mimarilerin konvansiyonel yaklaşımlara göre üstün performans sergilemelerine rağmen hala karşılaştıkları sorunlar bulunmaktadır. Önerilen modellerin derin öğrenme mimarilerinin sahip oldukları handikapların üstesinden gelinmesi yönünde olduğu anlaşılmaktadır. Ayrıca, özellik çıkarıcı olarak görev yapan kodlayıcı blokta kullanılan omurga mimarilerin etkinliklerinin araştırıldığı görülmektedir. Söz konusu motivasyon ile bu çalışmada, yapı yüzeylerindeki çatlakların semantik segmentasyonunda farklı kodlayıcı omurgalarının etkinlikleri irdelenmiştir. DeepLabv3+ modeli ile çatlakların optimum seviye segmente edilebilmesi için en uygun mimarinin belirlenmesi amaçlanmıştır. Farklı omurga yapılarının segmentasyon üzerinde etkinliklerinin daha genelleyici ve güvenilir olarak belirlenmesi için çatlak dağılımı ve morfolojisinin farklı olduğu veri setleri kullanılmıştır. Yürütülen işlemler akabinde, bu çalışmanın literature sunması muhtemel katkılar aşağıdaki gibi özetlenmiştir.

- Taranmış görüntülerden otomatik ve hassas bir şekilde çatlakların segmentasyonu için derin segmentasyon ağı olan DeepLabV3+ mimarisinin performansı kapsamlı olarak sunulmuştur.
- Kodlayıcı bloğun esas bileşenini oluşturan omurga mimarilerinin özellik çıkarma kabiliyetleri ile birlikte segmentasyon çerçevesi için eğitilebilir parametre açısından verimlilikleri de irdelenmiştir.
- Alçak ve yüksek seviyeli özelliklerin segmentasyon performansı üzerindeki etkinlikleri, nicel bulgular ile birlikte tartışılmıştır.
- Önerilen modellerin performanslarını güvenilir bir şekilde değerlendirmek için 10 katlı çapraz doğrulama yöntemi uygulanmıştır. Eğitim, doğrulama ve test verilerinin farklı olduğu her bir segmentasyon katında elde edilen sonuçların ortalaması alınarak, nihai sonuçlar belirlenmiştir.



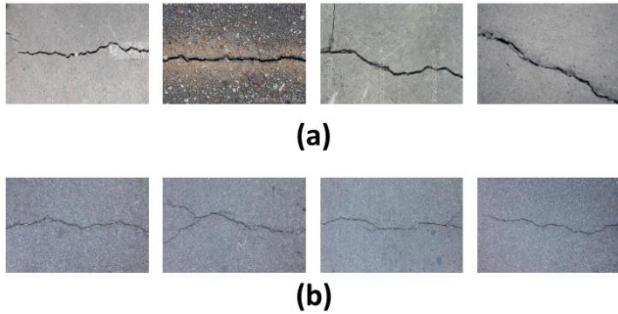
Şekil 1. Önerilen genel segmentasyon çerçevesi

Materyal ve Metot

Beton yapılarıdaki çatlakların piksel düzeyinde segmentasyonu için özellik çıkarıcı olarak farklı omurga mimarilerine ilişkin etkinliğin araştırıldığı genel segmentasyon çerçevesi Şekil 1’de tasvir edilmiştir. Bu bölümde, semantik segmentasyonda kullanılan veri setleri, DeepLabV3+ tabanlı genel çerçeve ve özellik çıkarımında etkinliği değerlendirilen evrişimsel sinir ağları (ESA) mimarileri (*ResNet-18*, *ResNet-50*, *MobileNetV2*, *Xception* ve *Inception-ResNet*) detaylandırılmıştır.

Veri seti

Bu çalışmada, önerilen modellerin etkinliklerinin test edilmesi için erişime açık DeepCrack ve CrackForest veri setleri kullanılmıştır. Veri setlerinin erişim adresleri sırasıyla kaynak [11], [30]’da paylaşılmıştır. Şekil 2’de her iki veri setine ait bazı örnek görüntüler gösterilmiştir.



Şekil 2. Veri setlerinden örnekler (a): DeepCrack, (b):CrackForest

(DeepCrack)

DeepCrack veri seti 384×544 boyutunda 537 RGB formatında görüntüden oluşmaktadır [11]. Veri seti beton ve asfalt gibi yüzeylerde oluşabilecek çatlaklara yönelik çeşitli ölçeklerdeki görüntüler kullanılması ile karakterize edilmiştir. Nispeten geniş hatlara sahip çatlaklar uzmanlarca işaretlenerek gerçek referans segmentasyon haritaları sağlanmıştır.

(CrackForest)

CrackForest veri seti 480×320 boyutunda toplam 118 yüzey görüntüsünden oluşmaktadır [30]. Çatlak içeren görüntülerin gerçek referans segmentasyon haritaları uzmanlar tarafından hazırlanarak, RGB formatında literatüre sunulmuştur.

DeepLabV3+

DeepLabV3+ kodlayıcı ve kod çözücü modüllerinden oluşan popüler bir semantik segmentasyon mimarisidir [23]. Kodlayıcı modülü temel olarak ham görüntülerden aşamalı olarak alt örneklerin alındığı ve bilgilerin ayıklandığı bölümdür. Kod çözücü modülü ise kodlayıcıdan gelen küçük boyutlu özellik haritalarından orijinal görüntü boyutunda segmentasyon haritalarının üretildiği bloktur [31].

Kodlayıcı bölümünde standart evrişim yerine derinlemesine ayrışabilir evrişimler kullanılarak uzamsal çözünürlük korunur [32]. Uzamsal çözünürlüğün korunduğu özellik haritaları atrous uzamsal piramit havuzlama (ASPP) blok ile birleştirilir. Böylelikle yüksek seviyeli özellik haritaları elde edilir. Kod çözücü blokta ise omurga mimari tarafından sağlanan alçak seviyeli özellikler ile yüksek seviyeli özellikler yukarı örneklenerek segmentasyon haritaları elde edilir.

Kodlayıcı Modülü

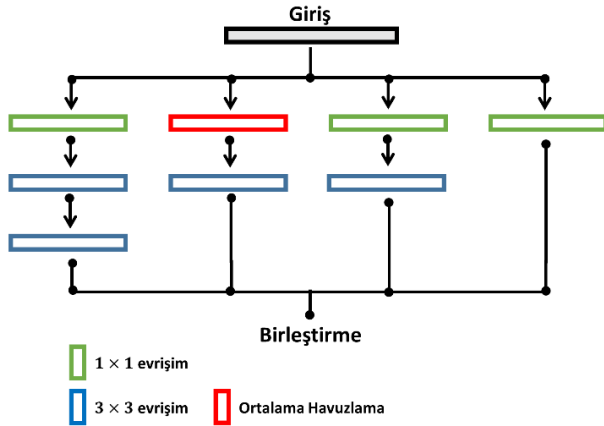
DeepLabV3+ segmentasyon çerçevesinde kodlayıcı modülü bir ESA mimarisinin uyarlanması ile gerçekleşir [2], [12]. Omurga mimari olarak uyarlanan ESA modelleri, görüntü sınıflandırma için tasarlandıklarından, güçlü özellik çıkarma kabiliyeti sunabilmektedir. Ancak, ESA mimarilerinin özellik çıkarımında üstün avantajlarına rağmen, probleme özgü sınırlı veri seti üzerinde eğitilmeleri çeşitli zorlukları da beraberinde getirir [3]. Veri seti boyutundaki sınırlılık nedeniyle iyi bir genelleştirme yapılamaması, arzu edilen ayırt edici semantik bilgilerin ayıklanmasını da olumsuz etkileyebilir. Bu kapsamda, transfer öğrenimi, söz konusu handikapların üstesinden gelinmesi için başvuru başlıca yaklaşımlardan biridir. Transfer öğrenimi ile ImageNet Büyük Ölçekli Görsel Tanıma Yarışması (ILSVRC) [33] gibi büyük ölçekli sınıflandırma problemleri üzerinde önceden eğitilmiş ESA modellerinin kodlayıcı bölümünde omurga yapı olarak kullanılması, etkin semantik bilgiler sunabilir.

DeepLabV3+ çerçevesi, en iyi semantik segmentasyon ağlarından biri olmasına rağmen, hesaplama etkinliği ağır olumsuz yönü olarak ön plana çıkmaktadır [2]. Ağın hesaplama maliyeti ve segmentasyon doğruluğu, omurga mimari ile doğrudan etkilenebilmektedir. Bu çalışmada yapılarda meydana gelen çatlakların segmentasyonunda, kodlayıcı bölümünde kullanılan farklı ESA modellerinin etkinliği değerlendirilmiştir. Bu kapsamda, ResNet, MobileNetv2, Xception ve Inception-ResNet derin mimarileri kullanılmıştır.

(ResNet)

ResNet büyük ölçekli görüntü sınıflandırma problemlerinde yüksek doğruluk sunan bir ESA mimarisi olup, He ve ark. [34] tarafından önerilmiştir. ResNet mimarisi evrişimsel ağlarda artan derinlikle birlikte kaybolan veya patlayan gradyan probleminin üstesinden gelinmesi için önerilmiştir. Bu kapsamda, atlamalı veya artık blok teniği takdim edilmiştir. Böylelikle, daha derin mimarilerin daha yüksek performansla eğitilmesi mümkün kılınabilmektedir [34].

Farklı varyantları bulunmasında rağmen, ResNet mimarisi genel olarak beş evrişim bloğundan oluşmaktadır. Artık bloklar evrişim bloklarının temelini oluşturmakta, evrişimsel katman sayısı her bir blokta farklılık göstermektedir [35]. Bu çalışmada ResNet-18 ve ResNet-50 mimarileri kullanılmıştır.



Şekil 3. Inception modülünün kanonik gösterimi

(Xception)

Xception modeli, derin mimarilerdeki Inception modülünün sıradan evrişim ile derinlemesine evrişim arasında yeniden dizayn edilerek türetildiği bir evrişimsel sınır ağıdır [36]. Şekil 3'te bir Inception modülünün kanonik gösterimi verilmiştir. Inception modülünde orijinal girişi sıkıştırmak için 1×1 boyutunda evrişim filtreleri (pointwise convolution) kullanmıştır. Önerilen ekstrem (extreme) Inception modülü ile Xception modelinde bu akış ters olarak işlenmiştir. Başka bir ifade ile pointwise evrişim uygulandıktan sonra kanal bazlı uzamsal evrişim işlemleri uygulanmıştır. Yeniden organize edilen Inception modülü ile giriş ve çıkışlarda küçük ölçekli farklılıklar ortaya çıkarak, söz konusu modifikasyon ile birlikte sınıflandırma probleminde daha üstün bir performans sağlanmıştır.

Xception mimarisi artık bağlantısı olan modüller halinde organize edilen 36 evrişim katmanından oluşur. Evrişim katmanları ağır özellik çıkarma yükünü üstlenmektedir. Batch normalizasyon işlemi ise tüm evrişimsel ve ayrılabilir evrişim katmanların akabinde uygulanır [37]. Üstün özellik çıkarımı yeteneği ile Xception modeli semantik segmentasyon modellerindeki kodlayıcı bloklar için omurga yapı olarak seçilmesinde başlıca motivasyon kaynağı olmuştur.

(MobileNet)

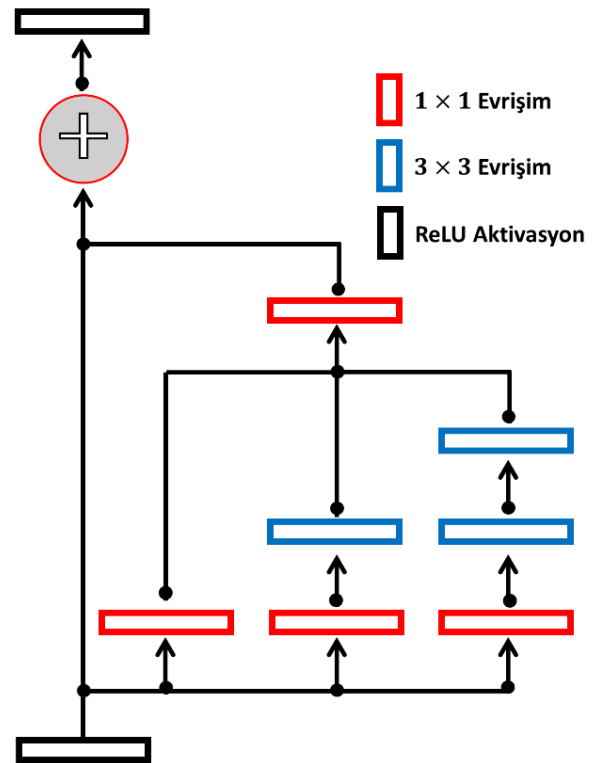
MobileNet esas olarak mobil ve gömülü görsel uygulamalar için Howard ve ark. [38] tarafından önerilmiştir. Oluşturulan mimari ile sınıflandırma problemlerinde doğruluk ile veri işleme süresi arasında verimli bir şekilde denge kurulması amaçlanmıştır. Ağ, ilk olarak Inception model ile takdim edilen derinlemesine ayrışabilir evrişimler üzerine inşa edilmiştir. Bu hafif mimaride standart evrişimler derinlemesine evrişim ve 1×1 boyutunda noktasal evrişimler olarak çarpanlarına ayrıştırılmıştır. Derinlemesine evrişim her bir kanal için ayrı ayrı tek bir filtre olarak uygulanırken, noktasal evrişim ise derinlemesine evrişimden gelen çıktıları bir araya getirmek için uygulanmıştır. Her evrişim işlemi akabinde batch normalizasyon ve ReLU aktivasyon işlemi uygulanmıştır.

Derinlemesine ayrışabilir evrişimlere dayanan MobileNet, yüksek verimlilikte işleme hızı ile birlikte rekabetçi bir

sınıflandırma doğruluğu sunabilmiştir. Bu çalışmada, transfer öğrenimin kodlayıcı blokta etkinliğinin test edilmesinde MobileNet'in güncellenmiş versiyonu olan MobileNetv2 kullanılmıştır [39]. Yeni versiyonda evrişim katmanların yapısında hafif değişimler sunmuştur. MobileNetv2'de ilk katman ReLU aktivasyon ile birlikte 1×1 evrişimdir ve ikinci katman derinlemesine evrişim işlemi uygulanmıştır. Son olarak, noktasal 1×1 boyutlu evrişim herhangi bir doğrusallık olmadan uygulanmıştır. Ayrıca, artık katman yapısı ResNet mimarisinin yapısına göre değiştirilerek, büyük yüke gerek kalmadan derinlemesine evrişimin doğruluğun geliştirilmesi yardımcı olmuştur.

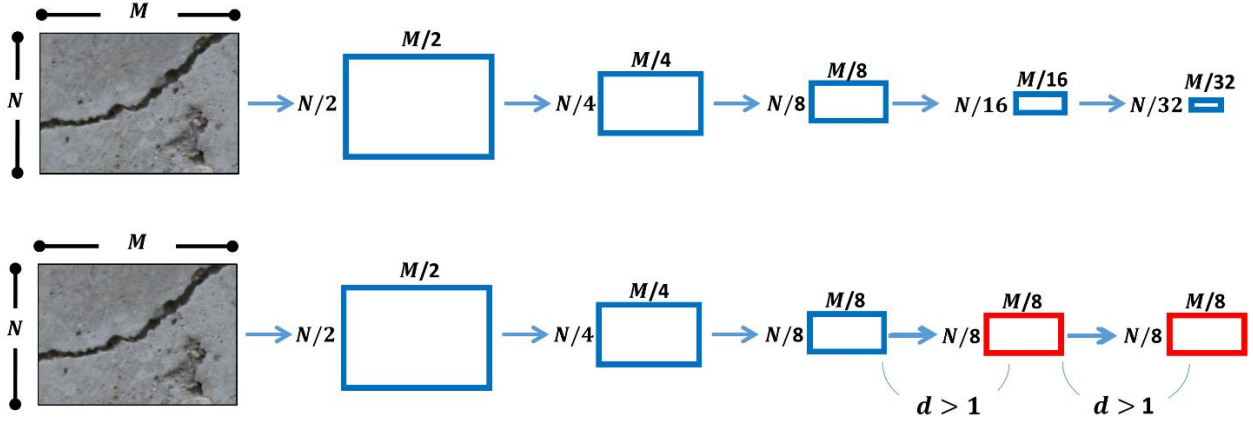
(Inception-ResNet)

Inception-ResNet modeli son yıllarda popüler olan artık bağlantılar ve Inception mimarisinin bir araya getirilme fikri ile önerilmiştir [40]. Artık bağlantılar daha derin mimarilerin daha etkin bir şekilde eğitilmesine olanak sağlayabilmiştir [41]. Inception modelinin son versiyonu olan Inception-v3'ün derin olması, derin mimaride filtre birleştirme aşamasının artık bağlantılar ile değiştirilmesi fikrini doğurmuştur. Bu değişim ile birlikte Inception mimarisinin hesaplama verimliliği korunur iken, ResNet modelinin tüm avantajlarından da istifade edilmesi olanaklı kılınmıştır. Şekil 4'te genel bir Inception-ResNet-V2 modeli gösterilmektedir.



Şekil 4. Inception-ResNet-v2 modeli/bloğu

Önerilen model ile artık mimarilerin Inception mimarisine eklemesi ile birlikte eğitim hızı ve performansının olumlu yönde etkilendiği tespit edilmiştir. ILSVRC 2102 veri seti üzerinde ResNet ve Inception-v3 modellerinin



Şekil 5. Genişletilmiş evrişim ile uzamsal çözünürlüğün korunması

sergilediklerinden daha üstün bir sınıflandırma performansı elde edilmiştir. Bu çalışmada, iki mimarinin kombinasyonundan türetilmiş Inception-ResNet-v2 modelinin segmentasyon için özellik çıkarma performansı irdelenmiştir.

Genişletilmiş (atrous) Evrişim ve ASPP

Omurga mimaride standart evrişim yerine kullanılan genişletilmiş (atrous) evrişim, aynı sayıda eğitilebilir parametre sayısı ile daha geniş alıcı alanlar sunabilmektedir [32]. Böylelikle, aynı görüntüden farklı ölçeklerde özelliklerin ayıklanılması mümkün kılınarak, çok ölçekli kavramsal bilgilerin elde edilmesi sağlanabilir [23]. Şekil 5'te uzamsal çözünürlüğün atrous evrişim ile korunması gösterilmektedir. Görüntülerden çıkarılan küçük boyutlu özellik haritaları özellikle sınıflandırmada problemlerinde başarılı bir performans sunarken, segmentasyon problemlerinde ise daha büyük özellik haritalarının kullanılması gerekir [31]. Bu kapsamda atrous evrişim operatörlerinin kullanılması, kodlayıcı bölümünde üretilen özellik haritalarındaki uzamsal bilgilerin korunması için önem arz etmektedir.

Kodçözücü Modül

Ham görüntülerden çıkış adımı (output stride) 16:1 olarak elde edilen semantik özelliklerden orijinal boyutta segmentasyon haritalarının üretilmesi kod çözücü blok tarafından sağlanır. Burada, kodlayıcı tarafından sağlanan özelliklerin yoğunlukları genişletilmiş evrişim ile kontrol edilebilir. Ancak genel olarak çıkış adımı 16 (Output-Stride) olarak seçilir [23].

Önerilen kod çözücü blokta ilk olarak kodlayıcıdan gelen özellikler faktör 4 ile çift doğrusal olarak yukarı örneklenir. Ardından, ön eğitilmiş ESA'dan oluşan omurga mimari tarafından sağlanan düşük seviyeli özellikler ile birleştirilir. Birleştirme işleminde düşük seviyeli özelliklerin aynı uzamsal çözünürlükte olması gerekmektedir. Omurga yapıdan gelen düşük seviyeli özellikler genellikle çok sayıda kanal içerdiğinden, kanal sayısını azaltmak için

başka bir 1×1 evrişim uygulanır. Son olarak, birleştirilen özellikler birkaç 3×3 filtre ile filtrelenip, yukarı örneklenerek segmentasyon işlemi tamamlanır.

Model Başarım Metrikleri

Semantik segmentasyon problemlerinde genellikle sınıflar arası dağılımın dengesiz olması kuvvetle muhtemeldir. Özellikle segmente edilecek ilgilen bölge, arka plana kıyasla daha az piksel ile temsil edilir. Dolayısıyla, önerilen modelin performansını güvenilir ve genelleyici olarak değerlendirmek için literatürde kabul edilen başlıca metriklere odaklanmak gerekir. Bu kapsamda, kesişimlerin birleşimi (IoU veya Jaccard indeksi) ve Dice benzerlik indeksi (DSC) sıklıkla tercih edilmektedir. IoU metriği, temel gerçek referans (ground-truth) ile model tarafından yapılan segmentasyon haritasının kesişim ve birleşimi temel alınarak hesaplanır. DSC metriği ise, kesişen piksel sayısının iki katı ile ground truth pikseller ve tahmin edilen piksellerin toplamı arasındaki oran olarak hesaplanır [42].

DSC ve IoU metriklerine ek olarak bu çalışmada özellikle segmentasyon ve sınıflandırma problemlerinde sıklıkla değerlendirilen doğruluk, kesinlik ve duyarlılık metrikleri de kullanılmıştır. Aşağıda söz konusu model başarım metriklerine ilişkin matematiksel eşitlikler verilmiştir.

$$DSC = \frac{2 \times TP}{2 \times TP + FP + FN} \quad (1)$$

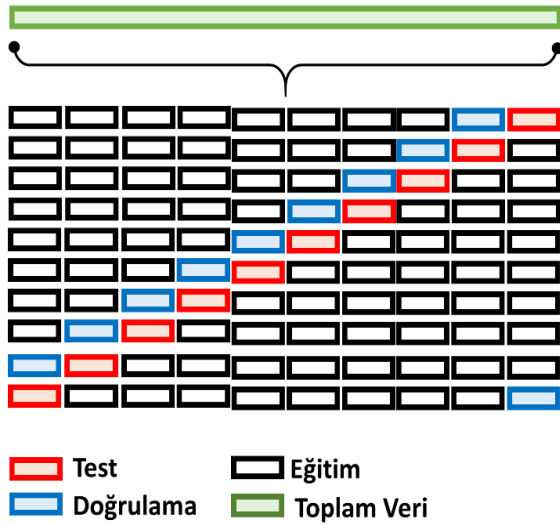
$$IoU = \frac{TP}{TP + FP + FN} \quad (2)$$

$$Duyarlılık = \frac{TP}{TP + FN} \quad (3)$$

$$Kesinlik = \frac{TP}{TP + FP} \quad (4)$$

$$Doğruluk = \frac{TP + TN}{TP + FP + TN + FN} \quad (5)$$

Değerlendirme metrikleri TP (Doğru Pozitif), FP (Yanlış Pozitif), TN (Doğru Negatif) ve FN (Yanlış Negatif) sayıları türünden hesaplanmaktadır.



Şekil 6. 10-katlı çapraz-doğrulama işlemi

Ağın Eğitimi ve Test Edilmesi

Önerilen segmentasyon modelleri ve tüm deneyler, MATLAB (MathWorks, Natick, Massachusetts) programlama ortamında, Intel i7-2.70 GHz işlemci, 8GB NVIDIA GeForce RTX grafik işleme birimi (GPU) ve 16 GB RAM ile donatılmış bir bilgisayarda uygulanmıştır. Farklı omurga mimarilerinden oluşan segmentasyon çerçeveleri uçtan uca ayrı bir şekilde eğitilmiştir. Veri setinin %80'i eğitim, %10'u doğrulama ve geri kalan %10'u ise test için kullanılmıştır. Önerilen modellerin performanslarını güvenilir bir şekilde değerlendirmek için 10 katlı çapraz doğrulama yöntemi uygulanmıştır. Eğitim, doğrulama ve test verilerinin farklı olduğu her bir segmentasyon katında elde edilen sonuçların ortalaması alınarak, nihai sonuçlar belirlenmiştir. Şekil 6'da 10 katlı çapraz doğrulama uygulaması gösterilmektedir. Şekil 6'da uygulanan veri dağılım metodolojisi her iki veri için de uygulanmıştır.

Tablo 1. DeepCrack veri setine ilişkin segmentasyon sonuçları

Yöntem	Sınıf	DSC		IoU		Doğruluk		Duyarlılık		Kesinlik	
		μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
DeepL. & Resnet-18	Çatlak	0.586	0.145	0.428	0.148	0.961	0.023	0.446	0.170	0.949	0.048
	Arkaplan	0.979	0.013	0.960	0.240	0.961	0.023	0.998	0.002	0.962	0.025
DeepL & Resnet-50	Çatlak	0.625	0.116	0.464	0.122	0.968	0.012	0.475	0.131	0.963	0.023
	Arkaplan	0.983	0.006	0.967	0.012	0.968	0.012	0.998	0.001	0.969	0.013
DeepL. & MobileNetV2	Çatlak	0.631	0.093	0.467	0.098	0.970	0.010	0.487	0.115	0.933	0.047
	Arkaplan	0.985	0.005	0.967	0.011	0.970	0.010	0.996	0.002	0.972	0.009
DeepL. & Xception	Çatlak	0.586	0.147	0.428	0.143	0.962	0.021	0.446	0.160	0.936	0.047
	Arkaplan	0.980	0.011	0.961	0.021	0.962	0.021	0.997	0.002	0.963	0.022
DeepL. & InceptionResnet	Çatlak	0.482	0.097	0.322	0.080	0.929	0.081	0.330	0.085	0.938	0.050
	Arkaplan	0.959	0.052	0.926	0.086	0.929	0.081	0.998	0.001	0.928	0.086

Not: DeepL=DeeplabV3+, μ = ortalama değeri, σ standart sapma

Hiperparmetre Seçimi

Bu çalışmada, derin öğrenme tabanlı öncelen segmentasyon ağları için eğitim aşamasında optimazson algoritması olarak Adam tekniği kullanılmıştır. Adam basit, hesaplama açısından verimli, hızlı ve parametreler açısından büyük olan problemler için verimli bir optimazyon algoritmasıdır [43]. Öğrenme oranı (learning rate) 0.001 seçilmiş ve her iki tam tur sonunda 0.4 oranında küçültülmüştür. Eğitim aşamasında her bir iterasyonda alınan örnek sayısı (mini-batch size), donanımsal kapasite doğrultusunda 8 olarak seçilmiştir. Tam tur sayısı (epoch) ise 10 olarak belirlemiştir. Eğitim aşamasındaki her bir parametrenin seçimi deneme ve yanılma yolu ile belirlenmiştir.

Deneyisel Bulgular

Bu bölümde, farklı ESA modellerinin omurga mimari olarak kullanıldıkları DeepLabV3+ çerçevelerin iki farklı erişime açık veri seti üzerinde segmentasyon başarımları paylaşılmıştır.

Tablo 1’de beş farklı ön eğitilmiş mimari kullanılarak 10-katlı çapraz-doğrulama sonucu elde edilen ortalama segmentasyon sonuçları DSC, IoU, duyarlılık, kesinlik ve doğruluk metrikleri türünden gösterilmiştir.

Farklı omurga mimarilerinin kullanıldığı segmentasyon modelleri için model başarımları metrikleri çatlak ve arkaplan sınıfları olmak üzere ayrı ayrı değerlendirilmiştir. Önerilen tüm modellerin genel olarak ham görüntülerden arka planı iyi bir şekilde ayırt edebildiği gözlemlenmiştir. Arka plan için segmentasyon performansı en yüksek olarak MobileNetV2 omurga mimari tarafından elde edilmiştir. MobileNetV2 ile segmentasyon sonuçları DSC, IoU ve doğruluk türünden sırasıyla %98.5, %96.7 ve %97 olarak elde edilmiştir. Arka plan için ise en düşük performansı Inception-ResNet omurga mimari tabanlı DeepLabV3+ segmentasyon çerçevesi sağlamıştır. Söz konusu mimari ile segmentasyon başarımları DSC, IoU ve doğruluk türünden sırasıyla %95.9, %92.6 ve %92.9 olarak elde edilmiştir.

Tablo 2. CrackForest veri setine ilişkin segmentasyon başarımları

Yöntem	Sınıf	DSC		IoU		Doğruluk		Duyarlılık		Kesinlik	
		μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
DeepL. & Resnet-18	Çatlak	0.399	0.047	0.250	0.036	0.952	0.018	0.280	0.057	0.773	0.167
	Arkaplan	0.975	0.010	0.952	0.019	0.952	0.018	0.992	0.012	0.959	0.013
DeepL & Resnet-50	Çatlak	0.410	0.057	0.259	0.046	0.953	0.021	0.284	0.053	0.787	0.159
	Arkaplan	0.975	0.011	0.952	0.021	0.953	0.021	0.992	0.012	0.958	0.014
DeepL. & MobileNetV2	Çatlak	0.419	0.044	0.266	0.036	0.958	0.013	0.312	0.089	0.750	0.173
	Arkaplan	0.978	0.007	0.958	0.014	0.958	0.013	0.991	0.013	0.966	0.011
DeepL. & Xception	Çatlak	0.406	0.073	0.257	0.059	0.953	0.019	0.291	0.081	0.780	0.169
	Arkaplan	0.975	0.010	0.952	0.019	0.953	0.019	0.992	0.014	0.959	0.016
DeepL. & InceptionResnet	Çatlak	0.361	0.094	0.226	0.067	0.929	0.046	0.237	0.071	0.859	0.106
	Arkaplan	0.961	0.265	0.927	0.047	0.929	0.046	0.995	0.005	0.931	0.046

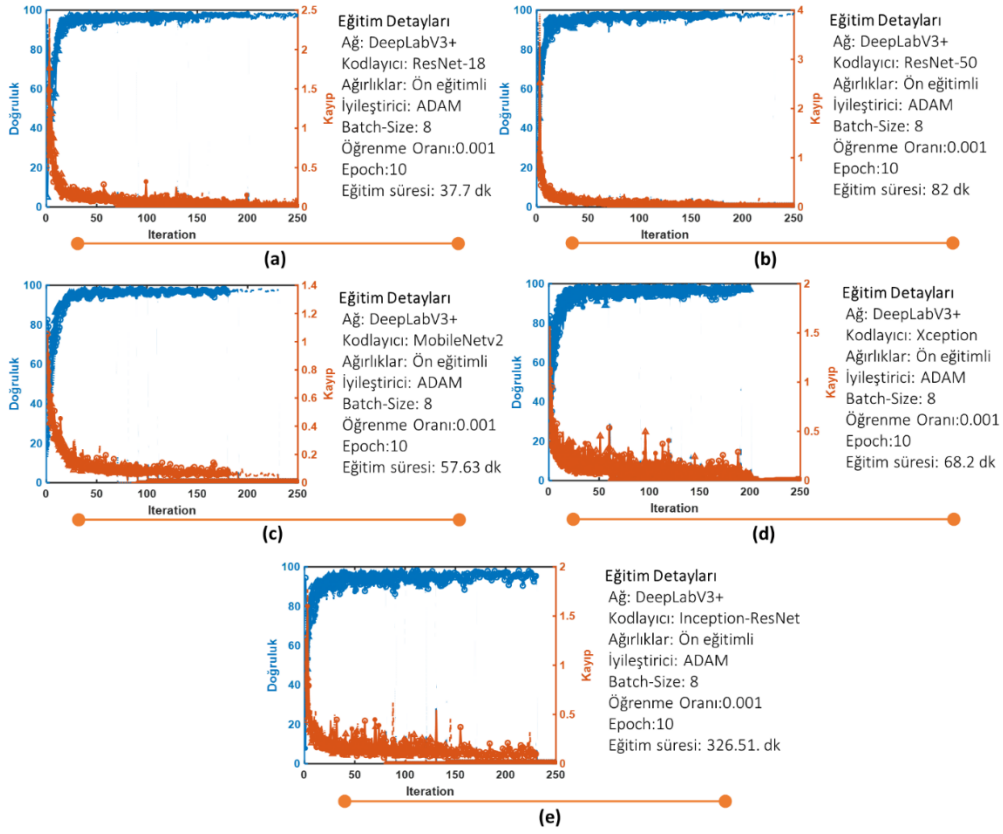
Not: DeepL=DeeplabV3+, μ = ortalama değer, σ standart sapma

Beton yüzeylere ait görüntülerden çatlakların segmentasyonu arka plana kıyasla genel olarak daha düşük bir performansla sağlanmıştır. Çatlak segmentasyonunda en yüksek performansı MobileNetV2 tabanlı segmentasyon çerçevesi sağlarken, en düşük performansı ise Inception+ResNet tabanlı segmentasyon çerçevesi sunmuştur. MobileNetV2 ile segmentasyon başarımları DSC, IoU ve doğruluk metriği türünden sırasıyla %63.1, %46.7 ve %97 olarak elde edilmiştir. MobileNetV2 omurga mimariye en yakın performans ResNet-50 modeli sunulmuş olup, segmentasyon başarımları sırasıyla %62.5, %46.4 ve %96.8 olarak elde edilmiştir. Çatlak segmentasyonuna ilişkin en düşük performans ise Inception+ResNet omurga mimari tarafından sunulmuştur. Inception+ResNet modelinin omurga mimari olarak kullanıldığı segmentasyon çerçevesi DSC, IoU ve doğruluk türünden %48.2, %32.2 ve %92.9 seviyesinde bir performans sergilemiştir. Tablo 2’de CrackForest veri setine ilişkin elde edilen segmentasyon başarımları gösterilmiştir.

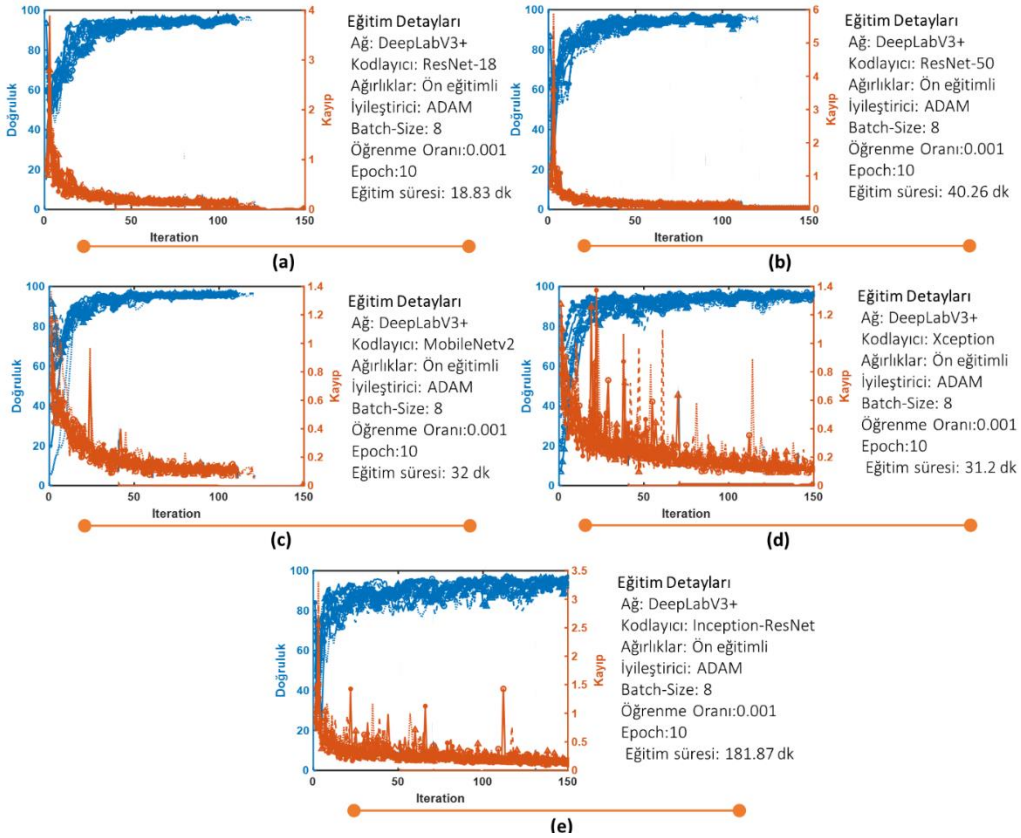
Tablo 2’de görüldüğü üzere CrackForest veri setine ilişkin elde edilen segmentasyon performansları DeepCack ile kıyaslandığında genel olarak daha düşük seviyede seyretmiştir. Ancak, en başarılı segmentasyon başarımlarını sunan omurga mimariler DeepCrack veri seti ile benzerlik göstermiştir. MobileNetV2 modelinin omurga mimari olarak kullanıldığı segmentasyon çerçevesi en yüksek başarımları sunmuştur. Arkaplan sınıfı için MobileNetV2 ile segmentasyon başarımları DSC, IoU, ve doğruluk metrikleri türünden sırasıyla %97.8, %95.8 ve %95.8 olarak elde edilmiştir. Beton yüzey görüntülerinde arkaplan sınıfına kıyasla çok az piksel ile temsil edilen çatlak sınıfına ilişkin segmentasyon başarımları daha düşük seviyede seyretmiştir. En yüksek başarımlar MobileNetV2 ile elde edilirken, en

düşük performans ise Inception-ResNet omurga mimarisi tarafından sunulmuştur. Her iki omurga mimari için segmentasyon başarımları DSC türünden sırasıyla %41.9 ve %36.1 olarak elde edilmiştir. Farklı segmentasyon mimarilerinin öğrenme sürecinin ilerlemesini gözlemlemek ve birbirleriyle karşılaştırmak için eğitim sırasında doğruluk ve kayıp eğrilerinin seyri incelenmiştir. Daha önce etiketlenmiş olan açıklamalı çatlaklar ile önerilen modeller tarafından tahmin edilen çatlaklar arasındaki mesafe böylelikle temsil edilebilmiştir. Özellikle eğitim kaybı eğrileri, eğitim veri setinin ağ tarafından yeterince öğrenilip öğrenilmediğini gösteren kritik göstergelerdir [12]. Şekil 7’de DeepCrack veri seti için farklı mimarilerin 10 katlı çapraz doğrulama ile sergiledikleri öğrenme kabiliyetlerine ilişkin eğitim doğruluğu ve eğitim kaybı eğrileri gösterilmiştir.

Şekil 7’de görüldüğü üzere tüm segmentasyon mimarileri için öğrenme işlemi ilk 50 iterasyonda önemli ölçüde tamamlanmıştır. Öğrenme aşamasında kayıp eğrisinde dalgalanmaların 50. iterasyondan sonra azaldığı ağların öğrenme sürecindeki son rötuşları gerçekleştirdiği gözlemlenmiştir. ResNet-18, ResNet-50 ve MobileNetV2’nin kodlayıcı bloktaki omurga olarak kullanıldığı DeepLabV3+ mimarileri için diğer omurga yapılarına kıyasla eğitim kaybının daha hızlı minimum seviyeye eriştiği rapor edilmiştir. Xception ve Inception-ResNet omurga yapıları için öğrenme aşamasında aşırı uyumlamanın daha erken gerçekleşmesinden ötürü öğrenme sürecinin erken sonlandığı çok sayıda segmentasyon katının olduğu gözlemlenmiştir. Şekil 8’de CrackForest veri seti için farklı mimarilerin 10 katlı çapraz doğrulama ile sergiledikleri öğrenme kabiliyetlerine ilişkin eğitim doğruluğu ve eğitim kaybı eğrileri gösterilmiştir.



Şekil 7. DeepCrack veri seri ilişkin eğitim ve kayıp fonksiyonları, (a) ResNet-18, (b) ResNet-50, (c) MobileNetV2, (d) Xception ve (e) Inception-ResNet



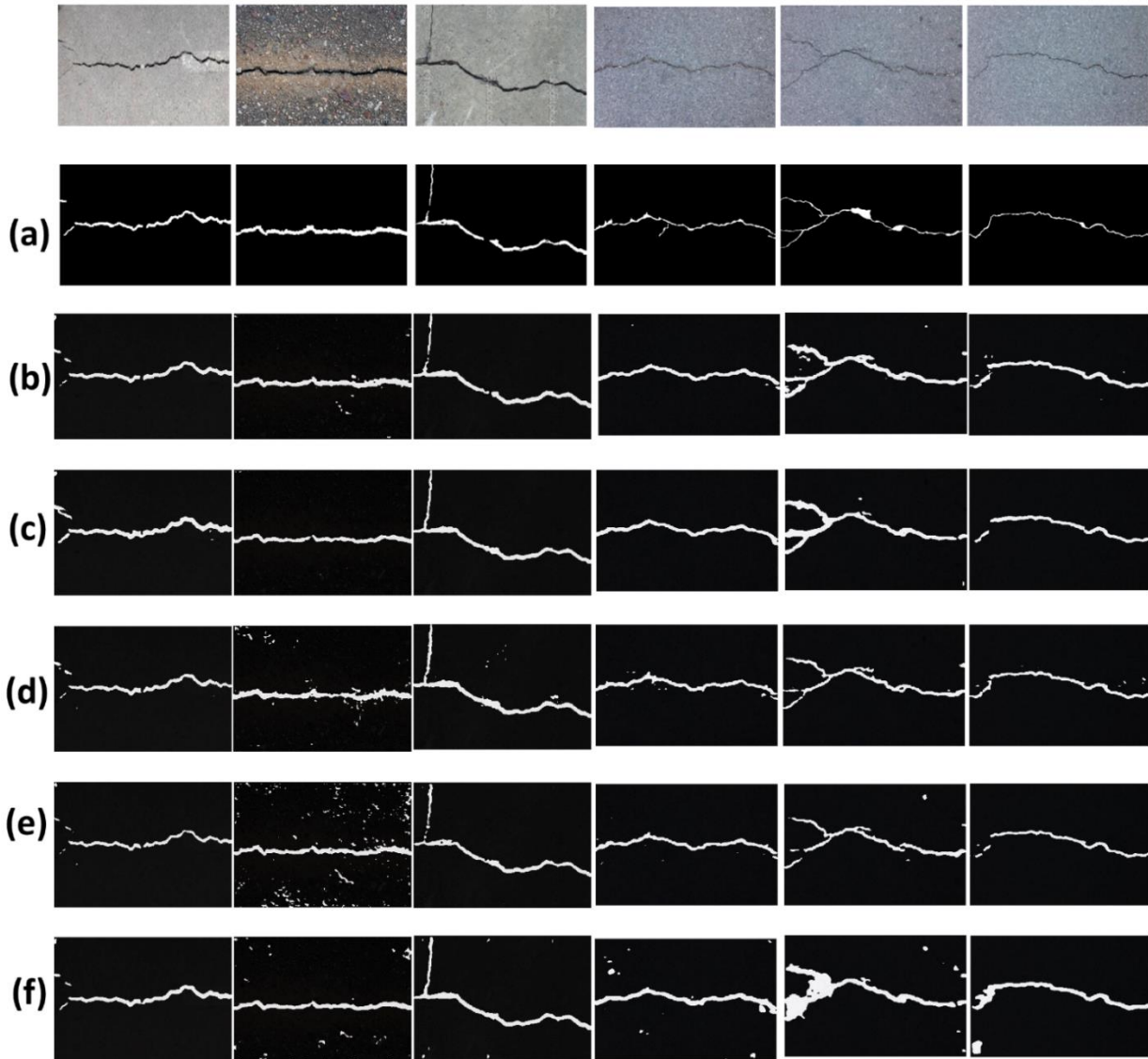
Şekil 8. CrackForest veri seri ilişkin eğitim ve kayıp fonksiyonları, (a) ResNet-18, (b) ResNet-50, (c) MobileNetV2, (d) Xception ve (e) Inception-ResNet

CrackForest veri seti üzerinde tüm segmentasyon modellerinin eğitim süreçlerinin daha düşük performansla seyrettiği gözlemlenmiştir. Eğitim kaybı eğrisinin 20. iterasyondan sonra hızla yakınsamaya başladığı gözlemlenmiştir. Eğitim doğruluğunun 50. iterasyondan sonra sabit değere yöneldiği gözlemlenmiştir. Özellikle Xception ve Inception-ResNet omurga yapı tabanlı segmentasyon mimarilerinin eğitim doğruluğu ve eğitim kaybı eğrilerinde yüksek dalgalanmaların olduğu tespit edilmiştir.

Farklı ön eğitilmiş ESA modellerine dayanan segmentasyon mimarilerinin sunmuş oldukları performansın güvenilir bir şekilde değerlendirilmesi için 10 katlı çapraz doğrulama yaklaşımı uygulanmıştır. Her bir segmentasyon katında farklı veri dağılımı senaryoları üretilerek, her bir görüntünün eğitim, doğrulama ve test grubunda yer alması

sağlanmıştır. Şekil 9’da farklı omurga mimarilere ilişkin segmentasyon çıktıların örnek görselleri verilmiştir.

Gerçek referans görüntüleri ile segmentasyon çıktıları kıyaslandığında farklı omurga yapılarının kullanıldığı DeepLabV3+ mimarilerinin genel olarak makul seviyede bir segmentasyon performansı sergilediği gözlemlenmiştir. Resnet-50 ve MobileNetV2 tabanlı mimarilerin daha net sonuçlar sunduğu görsel bulgulardan çıkarılabilmektedir. Ancak, Xception ve Inception-ResNet tabanlı segmentasyon mimarilerinin çok sayıda pikseli yanlış pozitif olarak etiketlediği tespit edilmiştir. DeepCrack veri setinde mevcut çatlakların nispeten daha kalın ve daha fazla piksel ile temsil edilmiş olmaları, önerilen modeller tarafından piksellerin yanlış negatif olarak etiketleme riskini de hafiflettiğine görsel segmentasyon çıktılarından ulaşılmıştır.



Şekil 9. Farklı mimarilere ilişkin elde edilen görsel segmentasyon çıktıları (a): Gerçek referans (Ground-Truth), (b) ResNet-18, (c): ResNet-50, (d): MobileNetv2, (e): Xception, (f): Inception-ResNetv2

Tartışma

Semantik segmentasyon, bilgisayarla görme alanında önemli bir araştırma konusu olarak öne çıkmakta ve özellikle son yıllarda beton yapılara ait görüntülerin analizinde sıklıkla kullanılmaktadır. Semantik segmentasyon mimarilerinde başarıyı artırmak için ön eğitilmiş ESA modellerin kullanılması popüler bir yaklaşımdır [25], [32], [37]. Elde edilen bulgular, ön eğitilmiş mimarilerin doğrudan segmentasyon performansını etkilediğini ortaya çıkarmıştır. Farklı ön eğitilmiş mimariler ile farklı düzeyde sınıflandırma performansının elde edilmesi, omurga yapı olarak kullanılan ESA mimarilerin önemini vurgulamıştır.

Yüksek hassasiyetli semantik segmentasyon modellerinin en büyük handikapları arasında hesaplama verimliliği gelmektedir [2]. Dolayısıyla, eğitilebilir parametre sayısı ve segmentasyon performansı açısından dengeli modellerin türetilmesi veya geliştirilmesi önem arz etmektedir. Elde edilen bulgular MobileNetV2 mimarisinin bu kapsamda en ideal omurga yapı olarak değerlendirilebileceğini göstermiştir. MobileNetV2 tabanlı DeepLabV3+ mimarisinin yaklaşık 6.7 milyon eğitilebilir parametre ile en yüksek performansı sağladığı rapor edilmiştir.

Esasen büyük ölçekli görüntü sınıflandırma problemleri için dizayn edilen ESA modelleri güçlü birer özellik çıkarıcıdır [7]. Bu çalışmada derin ve nistepen sığ ön eğitilmiş mimarilerin özellik çıkarmada performansları test edilmiştir. Derin kodlayıcı mimariler tarafından üretilen küçük özellik haritaları sınıflandırma problemleri için verimli olmalarına rağmen, küçük özellik haritaları segmentasyon için çeşitli dezavantajları beraberinde getirir [31]. Küçük özellik haritalarının çıkarılması, çatlak gibi küçük ölçekli bölgelere ilişkin uzamsal bilgilerin kaybolmasına yol açabilmektedir. Inception-Resnet mimarisinin diğer yapılara kıyasla daha derin olmasına rağmen daha düşük segmentasyon performansı sağlaması, büyük özellik haritalarının beton çatlakları için daha elverişli olduğunu göstermiştir.

Bu çalışmada önerilen modeller farklı iki veri seti üzerinde uygulanmış ve her iki veri seti için nispeten birbirinden uzak performanslar elde edilmiştir. Veri kümelerinde yer alan çatlakların geometri, boyut, aydınlatma ve gölgeleme açısından farklı özellikte olması model başarımlarını doğrudan etkilediği düşünülmektedir. Ayrıca, DeepCrack veri seti için elde edilen segmentasyon başarımının daha yüksek olmasının başlıca nedenlerinden biri de veri seti boyutunun daha büyük olmasına dayandırılmaktadır. Her iki veri setinin sınırlı sayıda örnek içeriyor olmasından ötürü performans düşüklüğünün üstesinden gelinmesi için literatürde sıklıkla veri artırımı yöntemlerine başvurulduğu gözlemlenmiştir.

Sonuç

Bu çalışmada, beton yüzeylerindeki çatlakların segmentasyonu için önerilen DeepLabV3+ mimarisinde farklı omurga mimarilerin etkinlikleri analiz edilmiştir. Kodlayıcı bloğun esas bileşenini oluşturan omurga mimarilerinin özellik çıkarma kabiliyetleri ile birlikte

segmentasyon çerçevesi için eğitilebilir parametre açısından verimlilikleri de irdelenmiştir. Bu kapsamda, Resnet-18, ResNet-50, MobileNetV2, Xception ve Inception-Resnet mimarilerin kodlayıcı blokta etkinlikleri, iki farklı veri seti üzerinde herhangi bir veri artırımı yöntemine başvurulmadan belli koşullar altında test edilmiştir. Her iki veri seti için de MobileNetV2 mimarisinin eğitilebilir parametre ve segmentasyon performansı açısından en başarılı ağ olduğu tespit edilmiştir. MobileNetV2 tabanlı önerilen DeepLabV3+ segmentasyon modeli yaklaşık 6.7 milyon parametre ile DeepCrack ve CrackForest veri setleri için DSC türünden sırasıyla %82 ve %70 oranında bir performans sergilemiştir. En düşük performans ise Inception-ResNet mimarisi tarafından sağlanmış olup, DeepCrack ve CrackForest veri setleri için segmentasyon başarımı DSC türünden %70 ile %80 olarak elde edilmiştir. Elde edilen bulgular neticesinde, kodlayıcı blokta kullanılan ESA modellerinin DeepLabV3+ mimarisinin performansını doğrudan etkilediği tespit edilmiştir. Ayrıca, kodlayıcı blokta kullanılan ESA mimarilerinde artan derinlik ve karmaşıklıkla birlikte performansın olumlu etkilenmediği gözlemlenmiştir. Omurga yapıdaki artan derinlik ve karmaşıklığın yüksek hesaplama maliyeti ile birlikte hantal bir eğitim sürecine neden olduğu tespit edilmiştir.

Kaynakça

- [1] Q. An *et al.*, "Segmentation of Concrete Cracks by Using Fractal Dimension and UHK-Net," *Fractal Fract.*, vol. 6, no. 2, pp. 1–18, 2022, doi: 10.3390/fractalfract6020095.
- [2] L. Song, H. Sun, J. Liu, Z. Yu, and C. Cui, "Automatic segmentation and quantification of global cracks in concrete structures based on deep learning," *Meas. J. Int. Meas. Confed.*, vol. 199, no. June, 2022, doi: 10.1016/j.measurement.2022.111550.
- [3] X. Han *et al.*, "Structural damage-causing concrete cracking detection based on a deep-learning method," *Constr. Build. Mater.*, vol. 337, no. 196, 2022, doi: 10.1016/j.conbuildmat.2022.127562.
- [4] Y. Bai, H. Sezen, and A. Yilmaz, "End-to-end deep learning methods for automated damage detection in extreme events at various scales," *Proc. - Int. Conf. Pattern Recognit.*, no. c, pp. 5736–5743, 2020, doi: 10.1109/ICPR48806.2021.9413041.
- [5] W. Wang, C. Su, G. Han, and H. Zhang, "A lightweight crack segmentation network based on knowledge distillation," vol. 76, no. May, 2023.
- [6] L. Yang, H. Huang, S. Kong, and Y. Liu, "A deep segmentation network for crack detection with progressive and hierarchical context fusion," *J. Build. Eng.*, vol. 75, no. May, 2023, doi: 10.1016/j.job.2023.106886.
- [7] J. König, M. D. Jenkins, M. Mannion, P. Barrie, and G. Morison, "Optimized deep encoder-decoder methods for crack segmentation," *Digit. Signal Process. A Rev. J.*, vol. 108, 2021, doi: 10.1016/j.dsp.2020.102907.
- [8] A. Mahgoub, A. Talab, Z. Huang, F. Xi, and L.

- Haiming, "Detection crack in image using Otsu method and multiple filtering in image processing techniques," *Opt. - Int. J. Light Electron Opt.*, pp. 1–4, 2015, doi: 10.1016/j.jileo.2015.09.147.
- [9] P. Chun, S. Izumi, and T. Yamane, "Automatic detection method of cracks from concrete surface imagery using two-step light gradient boosting machine," pp. 61–72, 2021, doi: 10.1111/mice.12564.
- [10] Q. Zou, Y. Cao, Q. Li, Q. Mao, and S. Wang, "CrackTree: Automatic crack detection from pavement images," *Pattern Recognit. Lett.*, vol. 33, no. 3, pp. 227–238, 2012, doi: 10.1016/j.patrec.2011.11.004.
- [11] Y. Liu, J. Yao, X. Lu, R. Xie, and L. Li, "DeepCrack: A deep hierarchical feature learning architecture for crack segmentation," *Neurocomputing*, vol. 338, pp. 139–153, 2019, doi: 10.1016/j.neucom.2019.01.036.
- [12] R. Pu, G. Ren, H. Li, W. Jiang, J. Zhang, and H. Qin, "Autonomous Concrete Crack Semantic Segmentation Using Deep Fully Convolutional Encoder-Decoder Network in Concrete Structures Inspection," *Buildings*, vol. 12, no. 11, 2022, doi: 10.3390/buildings12112019.
- [13] K. Makantasis, E. Protopapadakis, A. Doulamis, N. Doulamis, and C. Loupos, "Deep Convolutional Neural Networks for efficient vision based tunnel inspection," *Proc. - 2015 IEEE 11th Int. Conf. Intell. Comput. Commun. Process. ICCP 2015*, pp. 335–342, 2015, doi: 10.1109/ICCP.2015.7312681.
- [14] F. Bahreini and A. Hammad, "Dynamic graph CNN based semantic segmentation of concrete defects and as-inspected modeling," *Autom. Constr.*, vol. 159, no. November 2023, p. 105282, 2024, doi: 10.1016/j.autcon.2024.105282.
- [15] H. Li, H. Zhang, H. Zhu, K. Gao, H. Liang, and J. Yang, "Automatic crack detection on concrete and asphalt surfaces using semantic segmentation network with hierarchical Transformer," *Eng. Struct.*, vol. 307, no. February, p. 117903, 2024, doi: 10.1016/j.engstruct.2024.117903.
- [16] W. Qayyum, R. Ehtisham, A. Bahrami, C. Camp, J. Mir, and A. Ahmad, "Assessment of Convolutional Neural Network Pre-Trained Models for Detection and Orientation of Cracks," pp. 1–16, 2023.
- [17] Y. J. Cha, W. Choi, G. Suh, S. Mahmoudkhani, and O. Büyüköztürk, "Autonomous Structural Visual Inspection Using Region-Based Deep Learning for Detecting Multiple Damage Types," *Comput. Civ. Infrastruct. Eng.*, vol. 33, no. 9, pp. 731–747, 2018, doi: 10.1111/mice.12334.
- [18] F. Panella, A. Lipani, and J. Boehm, "Semantic segmentation of cracks: Data challenges and architecture," *Autom. Constr.*, vol. 135, no. December 2021, p. 104110, 2022, doi: 10.1016/j.autcon.2021.104110.
- [19] M. Jamshidi, M. El-Badry, and N. Nourian, "Improving Concrete Crack Segmentation Networks through CutMix Data Synthesis and Temporal Data Fusion," *Sensors*, vol. 23, no. 1, 2023, doi: 10.3390/s23010504.
- [20] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9351, pp. 234–241, 2015, doi: 10.1007/978-3-319-24574-4_28/COVER/.
- [21] V. Badrinarayanan, A. Kendall, and R. Cipolla, "U-Net: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 12, pp. 2481–2495, 2017.
- [22] E. Shelhamer, J. Long, and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 4, pp. 640–651, Apr. 2017, doi: 10.1109/TPAMI.2016.2572683.
- [23] L. C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 11211 LNCS, pp. 833–851, 2018, doi: 10.1007/978-3-030-01234-2_49.
- [24] Y. Zhang, X. Gao, and H. Zhang, "Deep Learning-Based Semantic Segmentation Methods for Pavement Cracks," 2023.
- [25] T. Lee, J. H. Kim, S. J. Lee, S. K. Ryu, and B. C. Joo, "Improvement of Concrete Crack Segmentation Performance Using Stacking Ensemble Learning," *Appl. Sci.*, vol. 13, no. 4, 2023, doi: 10.3390/app13042367.
- [26] Z. Al-Huda, B. Peng, R. N. A. Algburi, M. A. Alantari, R. AL-Jarazi, and D. Zhai, "A hybrid deep learning pavement crack semantic segmentation," *Eng. Appl. Artif. Intell.*, vol. 122, no. November 2022, p. 106142, 2023, doi: 10.1016/j.engappai.2023.106142.
- [27] A. N. Soni, "Crack Detection in buildings using convolutional neural Network," no. September, 2020.
- [28] A. Liu, W. Hua, J. Xu, Z. Yang, and J. Fu, "Concrete crack segmentation based on multi-dimensional structure information fusion-based network," *Constr. Build. Mater.*, vol. 414, no. August 2023, 2024, doi: 10.1016/j.conbuildmat.2024.134982.
- [29] H. Lang, Y. Yuan, J. Chen, S. Ding, J. J. Lu, and Y. Zhang, "Augmented Concrete Crack Segmentation: Learning Complete Representation to Defend Background Interference in Concrete Pavements," *IEEE Trans. Instrum. Meas.*, vol. 73, pp. 1–13, 2024, doi: 10.1109/TIM.2024.3378205.
- [30] Y. Shi, L. Cui, Z. Qi, F. Meng, and Z. Chen, "Automatic road crack detection using random structured forests," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 12, pp. 3434–3445, 2016, doi: 10.1109/TITS.2016.2552248.
- [31] B. Baheti, S. Innani, S. Gajre, and S. Talbar, "Semantic scene segmentation in unstructured environment with modified DeepLabV3+," *Pattern Recognit. Lett.*, vol. 138, pp. 223–229, 2020, doi: 10.1016/j.patrec.2020.09.014.

- 10.1016/j.patrec.2020.07.029.
- [32] T. Ahmad, V. Gharehbaghi, J. Li, C. Bennett, and R. Lequesne, "Crack segmentation in the wild using convolutional neural networks and bootstrapping," *Earthq. Eng. Resil.*, vol. 2, no. 3, pp. 348–363, 2023, doi: 10.1002/eer2.52.
- [33] Jia Deng, Wei Dong, R. Socher, Li-Jia Li, Kai Li, and Li Fei-Fei, "ImageNet: A large-scale hierarchical image database," pp. 248–255, 2009, doi: 10.1109/cvprw.2009.5206848.
- [34] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2016-Decem, pp. 770–778, 2016, doi: 10.1109/CVPR.2016.90.
- [35] B. Baheti, S. Gajre, and S. Talbar, "Semantic Scene Understanding in Unstructured Environment with Deep Convolutional Neural Network," *IEEE Reg. 10 Annu. Int. Conf. Proceedings/TENCON*, vol. 2019-Octob, pp. 790–795, 2019, doi: 10.1109/TENCON.2019.8929376.
- [36] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," *Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017*, vol. 2017-Janua, pp. 1800–1807, 2017, doi: 10.1109/CVPR.2017.195.
- [37] R. E. Philip, A. D. Andrushia, A. Nammalvar, B. G. A. Gurupatham, and K. Roy, "A Comparative Study on Crack Detection in Concrete Walls Using Transfer Learning Techniques," *J. Compos. Sci.*, vol. 7, no. 4, 2023, doi: 10.3390/jcs7040169.
- [38] A. G. Howard *et al.*, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," Apr. 2017, [Online]. Available: <http://arxiv.org/abs/1704.04861>.
- [39] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 4510–4520, 2018, doi: 10.1109/CVPR.2018.00474.
- [40] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2016-Decem, pp. 2818–2826, 2016, doi: 10.1109/CVPR.2016.308.
- [41] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 8691 LNCS, no. PART 3, pp. 346–361, 2014, doi: 10.1007/978-3-319-10578-9_23.
- [42] D. Tabernik, M. Šuc, and D. Skočaj, "Automated detection and segmentation of cracks in concrete surfaces using joined segmentation and classification deep neural network," *Constr. Build. Mater.*, vol. 408, no. September, p. 133582, 2023, doi: 10.1016/j.conbuildmat.2023.133582.
- [43] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," *3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc.*, pp. 1–15, 2015.