



# Dikkat mekanizmalarının bölütleme mimarilerine entegrasyonu ve meme lenf düğümü görüntülerinde uygulanması

## Integration of attention mechanisms into segmentation architectures and their application on breast lymph node images

Ersan YAZAN<sup>1\*</sup> , Muhammed Fatih TALU<sup>2</sup> 

<sup>1</sup>Bilgisayar Teknolojileri Bölümü, Besni Meslek Yüksekokulu, Adıyaman Üniversitesi, Adıyaman, Türkiye.

eyazan@adiyaman.edu.tr

<sup>2</sup>Bilgisayar Mühendisliği Bölümü, Mühendislik Fakültesi, İnönü Üniversitesi, Malatya, Türkiye.

fatihtalu@gmail.com

Geliş Tarihi/Received: 10.03.2022

Düzeltilme Tarihi/Revision: 27.05.2022

doi: 10.5505/pajes.2022.07838

Kabul Tarihi/Accepted: 03.08.2022

Araştırma Makalesi/Research Article

### Öz

Motorize mikroskopların yaygınlaşması, hastadan alınan dokunun otomatik taranarak tek bir büyük görsele aktarılması ve bölütlemeye özel derin/çekişmeli ağların üretilmesi gibi yenilikler bölütleme probleminde uzman etiketlemesine oldukça yakın çıktılar otomatik üretebilmek için umut vermektedir. Özellikle dikkat modüllerinin klasik 3D-UNet veya GAN mimarilerine entegrasyonu bölütleme performanslarının iyileştirildiği bilinmektedir. Bu çalışmada histopatolojik görüntü bölütleme probleminin çözümünde dört farklı dikkat modülü (DAF, DAF3D, DANet ve MSA) etkileri analiz edilmiştir. DAF ve DAF3D modüllerinde tekil (SLF) ve çoklu (MLF) katman özellikleri birlikte kullanılırken, DANet ve MSA modüllerinde pozisyon ve kanal dikkat olmak üzere iki farklı mekanizma kullanılmaktadır. Yapılan deneysel çalışmalar sonucunda, DAF3D dikkat modülünün bölütleme doğruluğunu en yüksek seviyeye çıkardığı (0.76 mIoU ve 0.89 PA) görülmüştür. Aynı zamanda yaklaşımlar arasında en düşük bölütleme maliyetine (1 görüntü için 0.156 saniye) sahip olan yöntem yine DAF3D olmuştur.

**Anahtar kelimeler:** Dikkat mekanizması, Histopatoloji, Bölütleme, DANet, DAF, DAF3D.

### Abstract

Innovations such as the widespread use of motorized microscopes, the automatic scanning of the tissue taken from the patient and transferring it to a single large image, and the production of deep/adversarial networks specific to segmentation have increased the hope of automatically producing outputs very close to expert labeling in the segmentation problem. Particularly, it is known that segmentation performances are improved by integrating attention modules into classical 3D-UNet or GAN architectures. In this study, the effects of four different attention modules (DAF, DAF3D, DANet and MSA) were analyzed in solving the histopathological image segmentation problem. While single (SLF) and multiple (MLF) layer features are used together in DAF and DAF3D modules, two different mechanisms, position attention module and channel attention module, are used in DANet and MSA modules. As a result of the experimental studies, it has been seen that the DAF3D attention module maximizes the segmentation accuracy (0.76 mIoU and 0.89 PA). At the same time, the method with the lowest segmentation cost (0.156 seconds for 1 image) among the approaches was again DAF3D.

**Keywords:** Attention mechanism, Histopathology, Segmentation, DANet, DAF, DAF3D.

## 1 Giriş

Histopatolojik görüntüler içerisinde kanserli dokuların tespiti ve lokalizasyonunda bilgisayar destekli sistemlerin kullanılması, işlemin daha hızlı gerçekleştirilmesi ve insan kaynaklı hataların önüne geçilmesi açısından önem arz etmektedir. Bu nedenle bu alanda birçok çalışma gerçekleştirilmiştir. Histopatolojik görüntülerin bölütlenmesinde gerçekleştirilen çalışmalar CNN tabanlı eğitim mimarilerini kullanmayan klasik yöntemler ve derin öğrenme tabanlı yöntemler olarak iki sınıfta incelenebilir. Kmeans [1], morphology [2], watershed [3], thresholding [4], Active Contour Models (ACM) [5] gibi yöntemler bölütleme alanında kullanılan klasik yöntemlerdendir.

Birçok alanda uygulanan derin öğrenme yöntemleri, bölütleme alanında da kullanılmıştır. Histopatolojik görüntülerde bölütleme ve sınıflandırma alanında da çok sayıda derin öğrenme yöntemi kullanılmış ve önerilmiştir. Klasik Evrişimsel Sinir Ağları (ESA) mimarilerinin [6]-[8] haricinde U-Net mimarisi [9]-[11] histopatolojik görüntülerin bölütlenmesinde

en yaygın kullanılan mimarilerdendir. Bunların haricinde kodlayıcı-kod çözücü mantığına dayalı SegNet [12] mimarisi de birçok çalışmada kullanılmıştır. Yaygın olarak kullanılan bir diğer mimari ise DeepLab [13], mimarisidir. DeepLab ilk olarak 2014 yılında geliştirilmiş ve 2017 yılında DeepLabV3+ [14] versiyonu ile son halini almıştır.

Bölütleme çalışmalarında kullanılan derin öğrenme mimarisinden bağımsız olarak, başarı oranının artırılması amacıyla farklı yöntemler önerilmiştir. Bu yöntemlerden bir tanesi de Dikkat Mekanizmasıdır. ESA'lar tasarımı nedeniyle yerel özelliklere odaklanırlar ve global bağlam bilgilerini yeterince kullanamazlar [15]. Dikkat mekanizmalarının temel amacı yerel bağlam bilgileri ile global bağlam bilgilerinin birlikte kullanılarak odaklanılan bölgenin vurgulanıp, arka planın baskınlaştırılması ile belirli bir bölgeye dikkatin çekilmesini sağlamaktır. Böylelikle görüntünün her bir pikselinin daha iyi bir şekilde anlaşılmasını sağlarlar.

Dikkat mekanizması hareket tanıma [16], doğal dil işleme [17], görüntü üretme [18], görüntü sınıflandırma [19] gibi farklı alanlarda kullanılmıştır. Görüntü bölütleme alanında da birçok

\*Yazışılan yazar/Corresponding author

çalışmada farklı dikkat mekanizmaları önerilmiştir. Wang ve diğ. 2018 yılında ultrason görüntülerinin bölütlemesi için Derin Dikkat Özellikleri (DAF) [20] yöntemini önermişler. Bu çalışmada ESA'nın sığ katmanlarında prostat dışı gürültüyü bastırmak ve derin katmanlarda daha fazla prostat ayrıntısını özelliklere dönüştürmek için farklı katmanlardan entegre edilen çok düzeyli özelliklerden seçici olarak yararlanmak için dikkat mekanizmasını kullanmışlar ve daha sonra 2019 yılında başarımı daha yüksek olan DAF3D [21] yöntemini önermişler. Zhao ve diğ. ise iki yönlü bilgi yayılımı yapan Noktasal Uzaysal Dikkat Ağı (Point-wise Spataial Attention Network - PSANet) [22] önermişlerdir. Mevcut pozisyonun tahminine yardımcı olmak için diğer pozisyonlardaki bilgiler kullanılırken, geçerli pozisyonadaki bilgiler de diğer pozisyonların tahminine yardımcı olmak için kullanılmıştır. Bu şekilde iki yönlü olarak toplanan bağlamsal bilgiler yerel özelliklerle birleştirilerek işlenmiş. Li ve diğ., gerçekleştirdiği Piramit Dikkat Ağı (PAN) [23] çalışmasında ise piksel etiketlemede daha önemli özellikleri elde etmek için dikkat mekanizması ve uzaysal piramit yapısı birleştirilmiştir. Bunun için ağa Özellik Piramit Dikkati (Feature Pyramid Attention-FPA) ve Global Dikkat Yukarı Örneklem (Global Attention Upsample-GAU) modülünü eklemişler. FPA modülünde giriş görüntüsünden daha iyi özellikler çıkarmak için 3 farklı ölçekte konvolüsyon uygulanmış ve bu konvolüsyon sonuçları birleştirilmiş. GAU modülünde ise düşük seviyeli özellikler ve yüksek seviyeli özellikler ayrı bir şekillerde işlenmiş ve sonrasında birleştirilmiştir. Böylelikle farklı ölçek haritalarını daha etkili bir şekilde dağıtarak düşük seviyeli özelliklere yüksek seviyeli özelliklerin rehberlik etmesi sağlanmıştır. Yu ve diğ. farklı yerel bölgelerin özelliklerine farklı önem ağırlıkları atamak için yığın yapısında bir dikkat ağı olan S2GA [24] yöntemini önermişler. Fu ve diğ. hem konum hem de kanal bağımlılıklarını modelleyen DANet [25] yöntemini önermişler. Bu çalışmada her konumdaki özelliği, tüm konumlardaki özelliklerin ağırlıklı toplamına göre seçici bir şekilde toplayan Konum Dikkat Modülü (Position Attention Module-PAM) ve ilişkili özellikleri tüm kanal haritaları arasında entegre ederek seçici olarak birbirine bağlı kanal haritalarını vurgulayan Kanal Dikkat Modülü (Channel Attention Module-CAM) önerilmiştir. Bu çalışmada elde edilen sonuçlar PAN [23] yöntemi ile kıyaslandığında farklı sınıflarda farklı başarımların olduğu görülmüştür. Bazı sınıflarda DANet yöntemi daha yüksek başarımlar gösterirken bazı sınıflarda ise, PAN yöntemi daha yüksek başarımlar göstermiştir. Wu ve diğ. konum ve kanal dikkat modüllerini içeren DAN [26] yöntemini önermişler. Çalışmada önerdikleri yöntem, anlamsal bilgileri etkili bir şekilde yakalayabilmek için DAP (Deformable Attention Pyramid) modülü ve uzaysal ve anlamsal özellikleri bağlamak için FAI (Fusing Attention Interim) modülünü içermektedir. Bir diğer konum ve kanal bazında dikkat mekanizması içeren ağ yapısı da Sinha ve Dolz tarafından önerilmiştir [27]. MSA olarak adlandırılan bu yöntemde çok ölçekli bir dikkat modülü önerilmiştir. Farklı modüller arasına eklenen kayıp fonksiyonu ile dikkat mekanizmasının ilgisiz bilgileri ihmal etmesi ve ilgili özellik ilişkilerini vurgulayarak görüntünün daha ayırt edici bölgelerine odaklanması sağlanmıştır. Huang ve diğ. ise zaman karmaşıklığını  $O(N)^2$ 'den  $O(N)$ 'e indirgeyen CCNet [28] yöntemini önermişler. Bu yöntemde özellik haritasındaki her konum (piksel) için her birinin yalnızca seyrek bağlantılara sahip olduğu iki ardışık çapraz dikkat modülü kullanılmıştır. Böylelikle her konum için çapraz dikkat modülü, bağlamsal bilgileri yatay ve dikey yönlerinde toplamıştır. İki çapraz dikkat modülünü seri olarak birleştirilmesiyle, her konumun, verilen görüntüdeki tüm piksellerden bağlamsal bilgi toplaması

sağlanmıştır. Nvidia firması çalışanlarından Tao ve diğ. da çok ölçekli hiyerarşik bir dikkat mekanizması önermişlerdir [29]. Bu yöntemde önerilen dikkat mekanizması, teorik olarak [30]'de önerilen yöntemle benzetmekle birlikte bu çalışmada; sabit bir ölçek kümesinin her biri için tüm dikkat maskelerini öğrenmek yerine, bitişik ölçekler arasında göreceli bir dikkat maskesinin öğrenilmesi sayesinde her ölçek için yoğun bir maskenin öğrenilmesi dezavantajını ortadan kaldırmıştır. Ayrıca çalışmada yaklaşık 4 kat daha verimli bellek kullanımı ve daha hızlı eğitim gerçekleştirildiği belirtilmiştir. Bir diğer dikkat mekanizması da Li ve diğ. önerdiği Doğrusal Dikkat Mekanizmasıdır [31]. Bu çalışmada Taylor genişlemesinin denklem üzerindeki birinci dereceden kestirimine dayalı ve karmaşıklığı  $O(N)^2$ 'den  $O(N)$ 'e düşüren doğrusal bir dikkat mekanizması önerilmiştir. Islam ve diğ., dikkat modülünü 3DUNet'e entegre ettikleri beyin tümörü bölütlemesi çalışmasında [32] dikkat modülünün tahmin başarımını arttırmadaki etkisini göstermişlerdir. Zhang ve diğ. ise Covid 19 CT görüntüleri üzerinde lezyon bölütlemesi gerçekleştirdikleri çalışmada [33] dikkat modülünü GAN mimarisi ile birlikte kullanmışlardır. Çalışma sonucunda elde ettikleri bulgulara göre dikkat modülü içeren ağı, dikkat modülü olmayan GAN ve DCGAN mimarilerine göre daha yüksek başarımlar elde ettiklerini belirtmişlerdir.

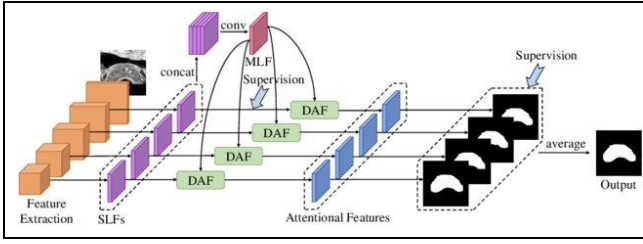
Literatürdeki çalışmalar incelendiğinde genel olarak bölütleme mimarilerine odaklandığı, bununla birlikte farklı dikkat mekanizmalarının bu mimarilere entegrasyonunun yapıldığı görülmektedir. Ancak dikkat mekanizmalarının performans katkısı karşılaştırmalı bir şekilde incelenmemiştir. Bu eksikliğin giderilmesi için bu çalışmada aynı veri kümesi üzerinde farklı dikkat mekanizmalarının (DAF, DAF3D, DANet ve MSA) bölütleme performansı üzerine etkisi karşılaştırmalı olarak incelenmesi gerçekleştirilmiştir.

## 2 Materyal ve metod

Bu çalışmada, son yıllarda önerilmiş yüksek başarımlar elde edilen ve birinci bölümde bahsedilmiş olan DAF, DAF3D, DANet ve MSA yöntemleri ile Camelyon 16 veri kümesinin [34] eğitim ve test işlemleri gerçekleştirilip yöntemlerin bölütlemedeki başarımları karşılaştırılmıştır. Bu bölümde yöntemlerin detaylı açıklamaları yer almaktadır.

### 2.1 Deep attentional features (DAF)

Transrektal ultrason (TRUS) görüntülerinden prostat bölütlemesi için gerçekleştirilen çalışmada çok ölçekli bir mimari önerilmiştir. Önerilen mimaride, ilk olarak ESA ile farklı çözünürlüklerde özellik haritaları çıkarılmıştır. Farklı bilgi seviyelerine sahip özellik haritaları elde edildikten sonra, farklı çözünürlükteki bu özellik haritaları, doğrusal entegrasyon ile orijinal giriş görüntüsünün dörtte biri oranında büyütülmüştür. Her bir katmandaki büyütülmüş özellik haritaları, tek katmanlı özellikler (SLF) olarak nitelendirilmiş ve bu SLF'ler, çok katmanlı özellikler (MLF) oluşturmak için konvolüsyon işlemleriyle birleştirilmiştir. Sonrasında Şekil 1'de görüldüğü gibi SLF ve MLF dikkat modülüne (DAF) giriş olarak verilmiştir. DAF modülünde gürültülerden arındırılmış, anlamsal ve detay bilgilerinin artırıldığı iyileştirilmiş bölütleme haritası elde edilmiş. Ardından tüm katmanlara uygulanan dikkat mekanizması ile her katmandaki derin dikkat özelliklerinden bölütleme haritaları elde edilmiştir.



Şekil 1. Bölütleme ağının mimarisi (DAF) [20].

Figure 1. Architecture of the segmentation network [20].

Son olarak, her bir katmandaki bölütleme haritalarının ortalaması alınarak çıkış görüntüsü elde edilmiştir.

Çalışmada sığ katmanlardaki özellik haritalarının prostatın ayrıntılı bilgilerini içerdiği, ancak aynı zamanda prostat olmayan bölgeleri de içerdiği ve derin katmanlardaki özellik haritalarının da prostatın konumunu belirtmek için anlamsal bilgileri yakalayabildiği belirtilmiştir. DAF modülü ile hem sığ katmanlardaki hem de derin katmanlardaki avantajlar kullanılmak istenmiştir. Şekil 2’de gösterilen DAF modülünde tek katmanlı özellik haritaları çok katmanlı özellik haritaları ile birleştirilerek  $F_x$  üretilmiş ve ardından Eşitlik 1’de gösterilen normalize edilmemiş dikkat ağırlığı  $W_x$  elde edilmiştir.

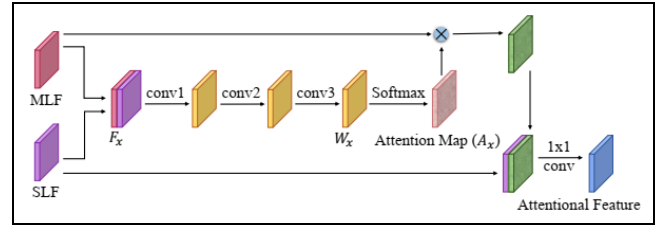
$$W_x = f_a(F_x; \theta) \quad (1)$$

Eşitlik 1’de  $\theta$ , üç evrişimli katman içeren  $f_a$  tarafından öğrenilen parametrelerdir.  $W_x$  elde edildikten sonra DAF modülü, Eşitlik 2’de gösterildiği gibi Softmax işlemi ile  $W_x$ ’i normalize ederek  $A$  dikkat haritasını hesaplamaktadır.

$$a_{i,j}^k = \frac{\exp(w_{i,j}^k)}{\sum_k \exp(w_{i,j}^k)} \quad (2)$$

Eşitlik 2’de gösterilen  $w_{i,j}^k$  normalize edilmemiş ağırlıklar ( $W_x$ ) içerisinde,  $k$  kanalındaki uzaysal konumu,  $a_{i,j}^k$  ise normalize edilmiş dikkat ağırlıklar ( $W_x$  içerisinde  $k$  kanalındaki uzaysal konumu temsil eder. Dikkat haritası elde edildikten sonra, MLF

ile noktasal çarpım yapılmış ve elde edilen yeni özellikler SLF ile birleştirilerek  $1 \times 1$  konvolüsyon uygulanmıştır. Böylelikle ilgili katman için dikkat özellikleri üretilmiştir.



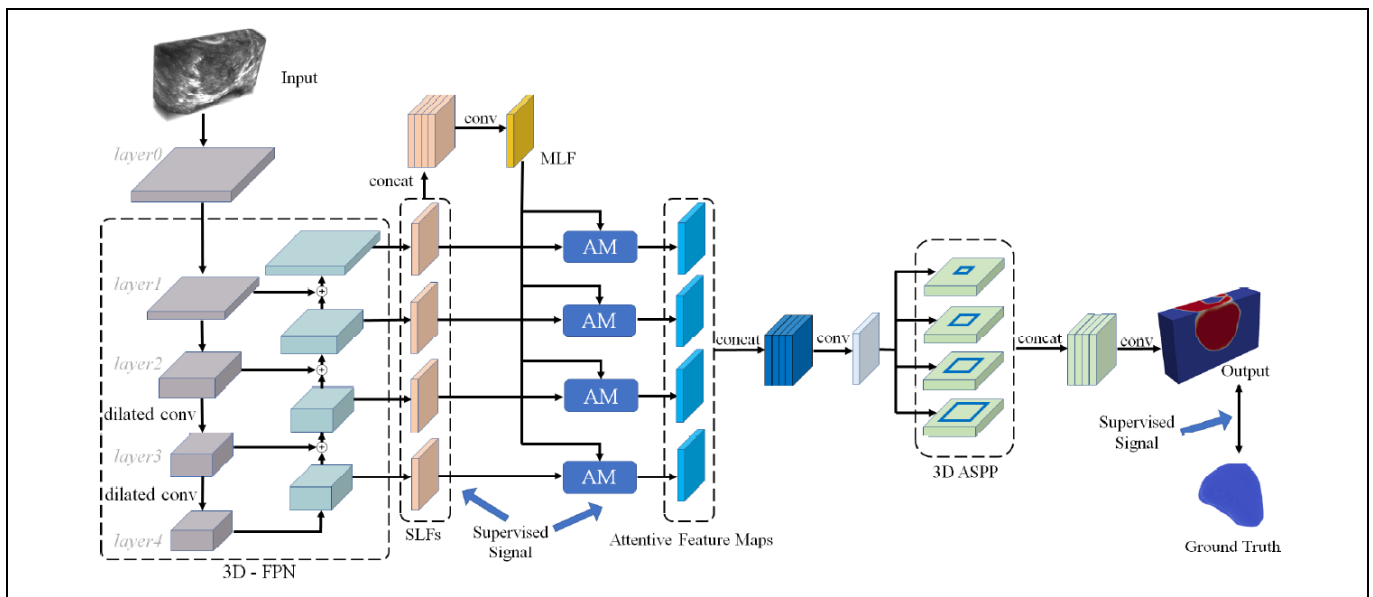
Şekil 2. DAF modülünün şematik gösterimi [20].

Figure 2. Schematic representation of the DAF module [20].

## 2.2 Üç boyutlu deep attentive features (DAF3D)

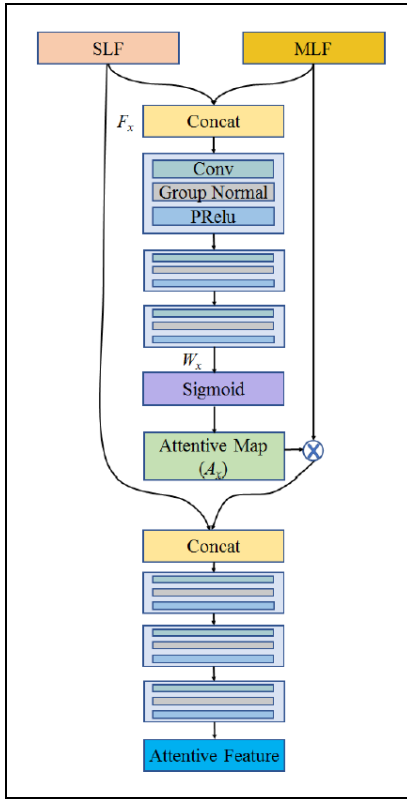
DAF yöntemiyle benzer yapıda olan bu yöntemde önerilen ağ mimarisinde Şekil 3’te gösterildiği gibi ESA tarafında bir piramit ağ (FPN) yapısı ile özellikler elde edilmiş, her katmana ait SLF, bir önceki katman bilgileri de kullanılarak oluşturulmuştur. Ağın çıkışında elde edilen dikkat özellik haritaları birleştirilerek ASPP (Atrous Uzaysal Piramit Ortaklama) işlemi uygulanmıştır.

Bu çalışmada önerilen dikkat modülünde Şekil 4’te görüldüğü gibi iki aşamalı bir işlem gerçekleştirilmiştir. İlk aşamada SLF ve MLF’ler birleştirilerek  $F_x$  üretilmiş ve ardından,  $F_x$  ağırlıklar (normalize edilmemiş) elde edilmiştir.  $W_x$  ağırlıklarını elde etmek için DAF yönteminde 3 tane peş peşe konvolüsyon işlemi uygulanırken bu çalışmada konvolüsyon, grup normalizasyonu ve PRelu işlemleri üç defa uygulanmıştır. Sonrasında da Sigmoid fonksiyonu ile normalize edilmiş  $A_x$  haritası elde edilmiştir. İkinci aşamada ise,  $A_x$  dikkat haritası MLF ile noktasal çarpıma tabi tutulmuş ve sonrasında SLF ile birleştirilmiştir. Elde edilen bu özellikler tekrardan 3 defa konvolüsyon, grup normalizasyonu ve PRelu işlemine tabi tutularak Dikkat özellikleri elde edilmiştir. Her iki aşamada da uygulanan 3 tane konvolüsyon işleminin ilk ikisi  $3 \times 3$ , sonuncusu ise  $1 \times 1$ ’lik maskeye sahiptir.



Şekil 3. DAF3D ağının genel görünümü [21].

Figure 3. Overview of the DAF3D network [21].



Şekil 4. DAF3D dikkat modülünün şematik gösterimi [21].  
Figure 4. Schematic representation of the DAF3D attention module [21].

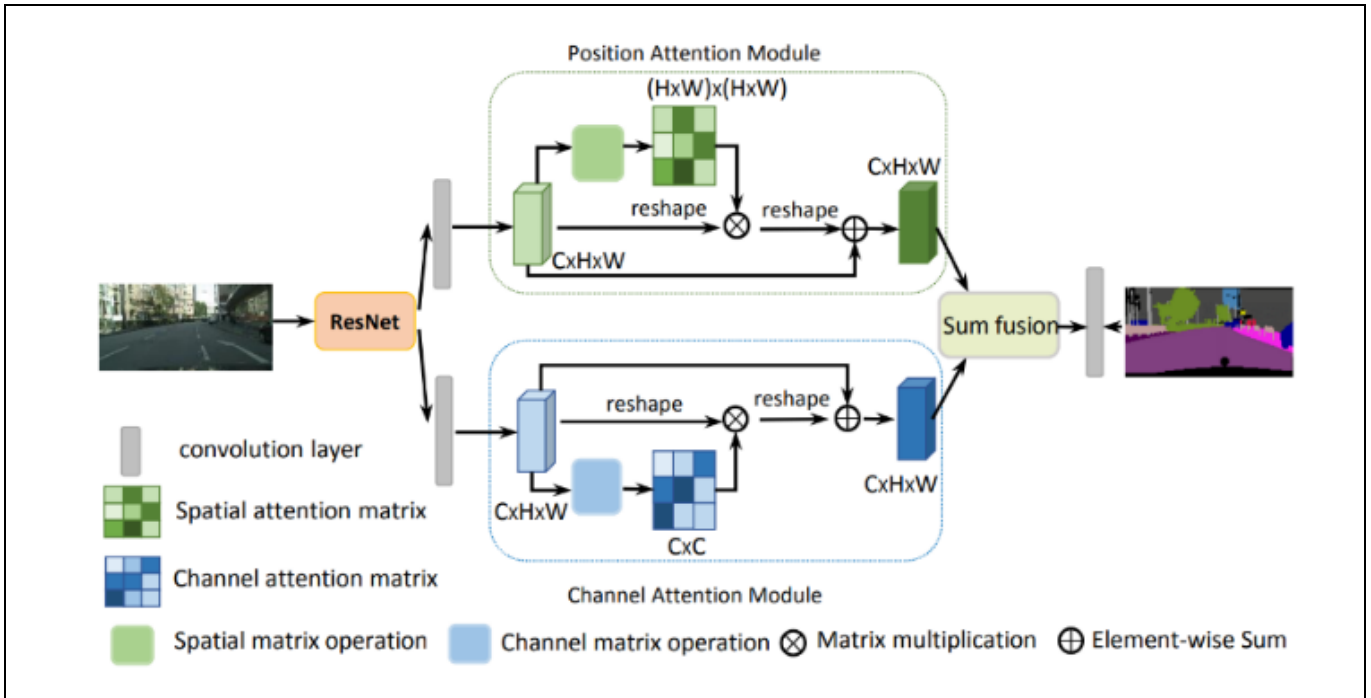
### 2.3 DANet

Sahne bölütlemesi için önerilen ve Cityscapes, PASCAL Context ve COCO Stuff veri setlerinin kullanıldığı bu çalışmada Konum Dikkat Modülü (KDM) ve Kanal Dikkat Modülü (KaDM) olmak üzere iki ayrı dikkat modülü birlikte kullanılmıştır. KDM ile her bir konumdaki özellik, tüm konumlardaki özelliklerin ağırlıklı toplamına göre seçici bir şekilde toplanarak mesafeden bağımsız olarak benzer özellikler birbirleriyle ilişkilendirilmiştir. KaDM ile de ilişkili özellikleri tüm kanal haritaları arasında entegre ederek seçici olarak birbirine bağlı kanal haritaları vurgulanmıştır.

Şekil 5'te gösterilen DANet mimarisinde, Konum Dikkat Modülü ile üç aşamada uzaysal uzun mesafeli bağlamsal bilginin yeni özellikleri elde edilmiştir. İlk aşamada, özelliklerin herhangi iki piksel arasındaki uzaysal ilişkiyi modelleyen bir uzaysal dikkat matrisi oluşturulmuştur. Ardından, dikkat matrisi ile orijinal özellikler arasında bir matris çarpımı gerçekleştirilmiştir. Üçüncü aşamada da, uzun mesafeli bağlamları yansıtan son gösterimleri elde etmek için, ikinci aşamada elde edilen çarpılmış sonuç matrisi ve orijinal özellikler üzerinde eleman bazlı bir toplama işlemi gerçekleştirilmiştir.

#### 2.3.1 Konum dikkat modülü (KDM)

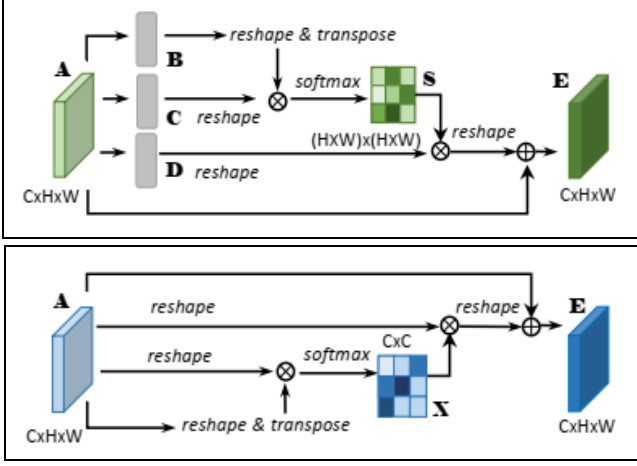
Bu modülde, Şekil 6'da gösterildiği gibi yerel bir özellik olan A girdisi ilk olarak bir konvolüsyon gerçekleştirilerek  $\{B, C, D\} \in R^{C \times H \times W}$  olacak şekilde üç yeni özellik olan B, C ve D oluşturulur. Sonrasında bu özellik haritaları,  $R^{C \times N}$  olacak şekilde yeniden şekillendiriliyor ( $N = H \times W$ ). Daha sonra her iki özellik matrisi çarpılarak uzaysal dikkat haritası S'yi  $S \in R^{N \times N}$  elde etmek için Softmax işlemine tabi tutulur. Eşitlik 3'te gösterilen bu işlemde  $s_{i,j}$  i. konumun j. konum üzerindeki etkisini ölçer.



Şekil 5. DANet ağının şematik gösterimi [25].

Figure 5. Schematic representation of the DANet network [25].

$$s_{ji} = \frac{\exp(B_i \cdot C_j)}{\sum_{i=1}^N \exp(B_i \cdot C_j)} \quad (3)$$



Şekil 6. Dikkat modüllerinin detaylı gösterimi [25]. (a): Konum dikkat modülü. (b): Kanal dikkat modülü.

Figure 6. Detailed representation of attention modules [25]. (a): Position attention module. (b): Channel attention module.

S dikkat haritası elde edildikten sonra D ile S'in transpozunu çarpılıp sonuç  $R^{C \times H \times W}$  olacak şekilde yeniden şekillendirilir. Son olarak elde edilen bu sonuç ölçek parametresi olan  $\alpha$  ile çarpılıp elde edilen çarpım girdi verisi olan A ile eleman bazlı toplanarak E çıktısı elde edilir. Bu işlemler de Eşitlik 4'te gösterilmiştir.

$$s_{ji} = \alpha \sum_{i=1}^N (s_{ji} D_i) + A_i \quad (4)$$

### 2.3.2 Kanal dikkat modülü

Kanal Dikkat Modülü kanallar arasındaki karşılıklı bağımlılıkları açıkça modellemek için önerilmiştir. Şekil 7'de görüldüğü gibi, KDM'den farklı olarak kanal dikkat haritası doğrudan orijinal özelliklerden hesaplanmaktadır. İlk olarak orijinal özellik haritası A, yeniden şekillendirilerek  $R^{C \times H \times W}$  boyutlarına getirilir ve transpozunu ile çarpılır. Sonrasında da kanal dikkat haritası X'i elde etmek için Eşitlik 5'te gösterilen softmax işlemi uygulanır ( $X \in R^{C \times C}$ ).

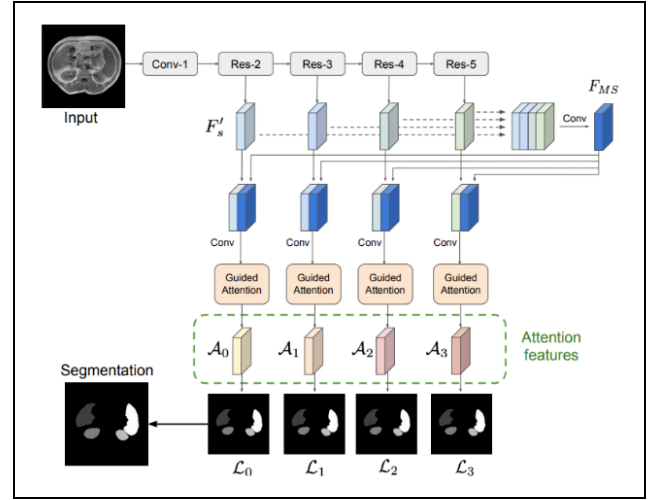
$$x_{ji} = \frac{\exp(A_i \cdot A_j)}{\sum_{i=1}^C \exp(A_i \cdot A_j)} \quad (5)$$

Eşitlik 5'te yer alan  $x_{i,j}$  i. kanalın j. kanal üzerindeki etkisini ifade eder. X dikkat haritası elde edildikten sonra transpozunu alınarak A ile çarpılır ve sonuç  $R^{C \times H \times W}$  olacak şekilde yeniden boyutlandırılır. Son olarak elde edilen matris ölçek parametresi  $\beta$  ile çarpılır ve çıktı matrisi E'yi elde etmek için Eşitlik 6'da gösterildiği gibi A ile toplanır.

$$E_j = \beta \sum_{i=1}^c (x_{ji} A_i) + A_j \quad (6)$$

## 2.4 Çok ölçekli dikkat ağı (MSA)

Abdominal organlar, kardiyovasküler yapılar ve beyin tümörlerini içeren veri kümelerinin bölütlemesinde çok ölçekli bir dikkat mekanizması kullanımının önerildiği bu çalışmada DANet ağında önerilen konum ve kanal dikkat modülleri kullanılmıştır. Bu dikkat modülleri sadece ESA'nın çıkışı yerine farklı katmanlara uygulanarak hem derin katmanlardaki hem de sıf katmanlardaki özelliklerden faydalanılması sağlanmıştır. Çalışmada dikkat modüllerinin farklı şekillerde kullanımları test edilmiş ve son olarak Şekil 7'deki mimari önerilmiştir.



Şekil 7. MSA ağının şematik gösterimi [27].

Figure 7. Schematic representation of the MSA network [27].

## 3 Bulgular ve tartışma

Bu çalışma kapsamında yapılan deneysel çalışmalar i7 işlemcili, NVIDIA Quadro M4000 ekran kartlı (8 GB GPU) ve 32 GB RAM'e sahip bilgisayar kullanılarak gerçekleştirilmiştir. Uygulama dili olarak Python seçilmiştir. Çalışmanın ikinci bölümde açıklamaları yapılan DAF, DAF3D, DANet ve MSA çalışmalarında anlatılan mimariler uygulanarak sonuçları karşılaştırılmıştır. Tüm yöntemlerde ESA mimarisi olarak ResNet50 [35] kullanılmıştır. Camelyon 16 veri kümesi üzerinde eğitim ve test işlemleri gerçekleştirilerek Pıksel Doğruluğu (Pixel Accuracy-PA) ve Ortalama Kesişimlerin Birleşimi (Mean Intersection over Union-mIoU) değerlendirme ölçütleri ile sonuçları değerlendirilmiştir.

Pıksel doğruluğu(PA), görüntüdeki doğru şekilde sınıflandırılmış pıksellerin yüzdesini belirten bir yöntemdir. Doğruluk oranı genel olarak her sınıf için ayrı ayrı veya tüm sınıflar için global olarak üretilebilir. Eşitlik 7'de belirtilen doğruluk hesabında doğru pozitif (TP) hedef maskedeki sınıfa ait olduğu doğru bir şekilde tahmin edilen bir pıkseli temsil ederken doğru negatif (TN) verilen sınıfa ait olmadığı doğru olarak tahmin edilen bir pıkseli temsil eder. Benzer şekilde yanlış pozitif (FP) hedef maskedeki sınıfa ait olduğu tahmin edilmesine rağmen o sınıfta olmayan bir pıkseli temsil ederken, yanlış negatif ise hedef maskedeki sınıfa olmadığı tahmin edilmesine rağmen o sınıfta olan bir pıkseli temsil eder.

$$PA = \frac{TP + TN}{TP + TN + FP + FN} \quad (7)$$

Jaccard endeksi olarak da adlandırılan kesişimlerin birleşimi yöntemi, hedef maske ile tahmin çıktısı arasındaki örtüşme yüzdesini ölçmek için kullanılan bir yöntemdir.

Basit bir şekilde hesaplaması olan yöntemde Eşitlik 8'de görüldüğü gibi, hedef ve tahmin maskeleri arasındaki ortak piksel sayısı, her iki maskede bulunan toplam piksel sayısına bölünerek hesaplanır.

$$IoU = \frac{\text{hedef} \cap \text{tahmin}}{\text{hedef} \cup \text{tahmin}} \quad (8)$$

Gerçekleştirilen çalışmada, 4330 adet eğitim verisi ve 433 adet test verisi kullanılmıştır. Optimizasyon yöntemi olarak Adam kullanılmış ve her bir yöntem için 10 ve 20 epoch olarak iki ayrı eğitim işlemi gerçekleştirilmiş ve elde edilen sonuçlar Tablo 1'de gösterilmiştir.

Tablo 1. Yöntemlerin bölütleme performansları.

Table 1. Segmentation performances of the methods.

	10 Epoch		20 Epoch	
	mIoU	PA	mIoU	PA
DAF	0.64	0.83	0.65	0.83
DAF3D	0.74	0.87	0.76	0.89
DANet	0.62	0.81	0.67	0.84
MSA	0.60	0.79	0.61	0.80

Tablo 1'de görüldüğü gibi 10 epoch eğitim sonucunda 0.74 mIoU ve 0.87 PA oranı, aynı veri seti ile 20 epoch eğitim sonucunda gerçekleştirilen test sonucunda da 0.76 mIoU ve 0.89 PA oranı ile DAF3D en yüksek performansı gösteren mimari olmuştur.

Bu mimarilerin çalışma zamanlarının kıyaslaması için 256x256x3 boyutundaki patolojik bir görüntüyü bölütleme zamanları ve kullandıkları hiperparametre sayıları incelenmiş ve elde edilen sonuçlar Tablo 2'de gösterilmiştir.

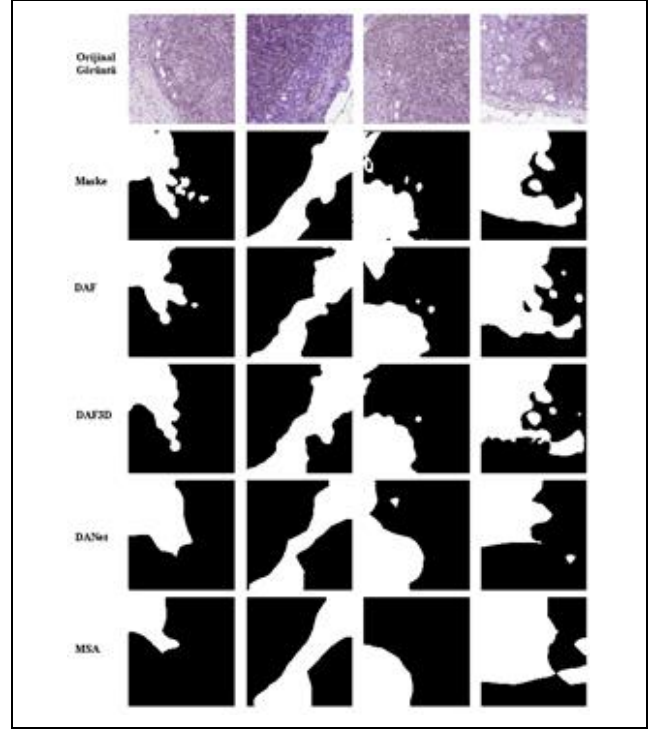
Tablo 2. Hiperparametre sayısı ve çalışma zamanı.

Table 2. Hyperparameter counts and running times.

	Hiperparametre sayısı (milyon)	Ortalama çalışma zamanı (sn.)
DAF	~43.488	0.255
DAF3D	~43.670	0.156
DANet	~66.058	0.182
MSA	~59.903	0.216

Tablo 2'de her bir sütunda parametre sayısı ve çalışma zamanları için en küçük değerler koyu olarak belirtilmiştir. Buna göre en az parametre sayısına sahip olan DAF mimarisidir. Bölütlemeye en iyi başarıya sahip olan DAF3D mimarisi ikinci olarak en az parametre sayısına sahip olmakla birlikte en düşük çalışma zamanına sahiptir. Deneysel çalışma sonucunda her bir mimarinin çıktıları kaydedilmiş ve bunlardan 4 tanesi Şekil 8'de gösterilmiştir. İlk satırda orijinal görüntü, ikinci satırda maske (referans görüntü) diğer satırlarda da mimarilerin çıktıları yer almaktadır.

Şekil 8'deki mimari bölütleme çıktıları incelendiğinde DAF3D odaklama yaklaşımının hedeflenen referans görüntüye en benzer çıktılar ürettiği gözle görülebilmektedir. DAF3D mimarisinin başarısı, tekil (SLF) ve çoklu (MLF) özniteliklerin birleşimini değerlendiren bir mimariye sahip olmasından kaynaklandığı görülmektedir. Elde edilen görsel sonuçların Çizelge 1'de verilen sayısal benzerlik sonuçlarıyla paralel oluşan deneysel çalışmaların tutarlılığını ortaya koymuştur.



Şekil 8. Girdi görüntüleri, maskeler ve çıktılar.

Figure 8. Input images, masks and outputs.

#### 4 Sonuçlar

Bu çalışmada histopatolojik görüntülerin bölütlenmesi probleminde dikkat mekanizmalarının etkisi analiz edilmiştir. Bu doğrultuda literatürde yaygın kullanılan dört farklı dikkat mekanizması (DAF, DAF3D, DANet ve MSA) Python programlama diliyle gerçekleştirildikten sonra mevcut bir bölütleme mimarisine entegrasyonları analiz edilmiştir. Deneysel çalışmalar sonucunda, DAF3D dikkat modülünün bölütleme doğruluğunu en yüksek seviyeye çıkardığı (0.76 mIoU ve 0.89 PA) görülmüştür. Bununla birlikte mimari parametre sayısı en düşük modül DAF olmasına karşılık DAF3D modülünün en düşük bölütleme maliyetine (1 görüntü için 0.156 sn.) sahip olduğu görülmüştür.

#### 5 Results

In this study, the effect of attention mechanisms on the problem of segmentation of histopathological images was analyzed. Accordingly, four different attention mechanisms (DAF, DAF3D, DANet and MSA), which are widely used in the literature, were implemented with the Python programming language, and their integration into an existing segmentation architecture was analyzed. As a result of the experimental studies, it was seen that the DAF3D attention module maximized the segmentation accuracy (0.76 mIoU and 0.89 PA). On the other hand, while DAF has the lowest number of architectural parameters, DAF3D module has the lowest segmentation cost (0.156 seconds for 1 image).

#### 6 Yazar katkı beyanı

Gerçekleştirilen çalışmada Ersan YAZAN literatür taraması, verilerin toplanması, analizlerin gerçekleştirilmesi ve elde edilen sonuçların değerlendirilmesi; Muhammed Fatih TALU fikrin ortaya konulması ve eleştirel incelenme, bulguların

incelenmesi, yazım denetimi ve makalenin içerik açısından kontrol edilmesi başlıklarında katkı sunmuşlardır.

## 7 Etik kurul onayı ve çıkar çatışması beyanı

Hazırlanan makalede etik kurul izni alınmasına gerek yoktur. Hazırlanan makalede herhangi bir kişi/kurum ile çıkar çatışması bulunmamaktadır.

## 8 Kaynaklar

- [1] Ray S, Turi RH. "Determination of number of clusters in k-means clustering and application in colour image segmentation". In *Proceedings of the 4<sup>th</sup> international conference on advances in pattern recognition and digital techniques*, Calcutta, India, 22-29 December 1999.
- [2] Park SH, Yun ID, Lee SU. "Color image segmentation based on 3-D clustering: morphological approach". *Pattern Recognition*, 31(8), 1061-1076, 1998.
- [3] Zhang X, Shan Y, Wei W, Zhu Z. "An Image Segmentation Method Based on Improved Watershed Algorithm". *2010 International Conference on Computational and Information Sciences*, Chengdu, China, 17-19 December 2010.
- [4] Tobias OJ, Seara R. "Image segmentation by histogram thresholding using fuzzy sets". *IEEE Transactions on Image Processing*, 11(12), 1457-1465, 2002.
- [5] Derraz F, Beladgham M, Khelif MH. "Application of active contour models in medical image segmentation". *International Conference on Information Technology: Coding and Computing*, Las Vegas, Nevada, USA, 5-7 April 2004.
- [6] Al-Milaji Z, Ersoy I, Hafiane A, Palaniappan K, Bunyak F. "Integrating segmentation with deep learning for enhanced classification of epithelial and stromal tissues in H&E images". *Pattern Recognition Letters*, 119, 214-221, 2019.
- [7] Pan X, Li L, Yang H, Liu Z, Yang J, Zhao L, Fan Y. "Accurate segmentation of nuclei in pathological images via sparse reconstruction and deep convolutional networks". *Neurocomputing*, 229, 88-99, 2017.
- [8] Bulten W, Hulsbergen-van de Kaa, CA, van der Laak J, Litjens GJ. "Automated segmentation of epithelial tissue in prostatectomy slides using deep learning". In *Medical Imaging 2018: Digital Pathology*, 10581, 219-225, 2018.
- [9] Matuszewski DJ, Sintorn IM. "Reducing the U-Net size for practical scenarios: Virus recognition in electron microscopy images". *Computer methods and programs in biomedicine*, 178, 31-39, 2019.
- [10] Kalapahar A, Silva-Rodríguez J, Colomer A, López-Mir F, Naranjo V. "Gleason Grading of Histology Prostate Images Through Semantic Segmentation via Residual U-Net". *2020 IEEE International Conference on Image Processing (ICIP)*, Abu Dhabi, United Arab Emirates, 25-28 October, 2020.
- [11] AskariHemmat M, Honari S, Rouhier L, Perone CS, Cohen-Adad J, Savaria Y, David JP. "U-net fixed-point quantization for medical image segmentation". *International Workshops, LABELS 2019, HAL-MICCAI 2019, and CuRIOUS 2019, Held in Conjunction with MICCAI 2019*, Shenzhen, China, 13-17 October, 2019.
- [12] Badrinarayanan V, Kendall A, Cipolla R. "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(12), 2481-2495, 2017
- [13] Chen LC, Papandreou G, Kokkinos I, Murphy K, Yuille AL. "Semantic image segmentation with deep convolutional nets and fully connected crfs". arXiv, 2016. <https://arxiv.org/pdf/1412.7062.pdf>
- [14] Chen LC, Papandreou G, Kokkinos I, Murphy K, Yuille AL. "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(4), 834-848, 2017.
- [15] Bozdağ Z. *Histopatolojik Görüntülerde Tümör Bölütlenmesi*. Doktora Tezi, İnönü Üniversitesi, Malatya, Türkiye, 2021.
- [16] Wang X, Girshick R, Gupta A, He K. "Non-local Neural Networks." *IEEE 2018 Computer Society Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, 18-23 June 2018.
- [17] Galassi A, Lippi M, Torroni P. "Attention in natural language processing". *IEEE Transactions on Neural Networks and Learning Systems*, 32(10), 4297-4308, 2020.
- [18] Zhang H, Goodfellow I, Metaxas D, Odena A. "Self-Attention Generative Adversarial Networks". *36<sup>th</sup> International Conference on Machine Learning, ICML*, Long Beach, California, USA, 9-15 June 2019.
- [19] Wang F, Jiang M, Qian C, Yang S, Li C, Zhang H, Tang X. "Residual attention network for image classification". *2017 IEEE conference on computer vision and pattern recognition*, Honolulu, Hawaii 21-26 July, 2017.
- [20] Wang Y, Deng Z, Hu X, Zhu L, Yang X, Xu X, Ni D. "Deep attentional features for prostate segmentation in ultrasound". In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Granada, Spain, 16-20 September, 2018.
- [21] Wang Y, Dou H, Hu X, Zhu L, Yang X, Xu M, Ni D. "Deep attentive features for prostate segmentation in 3D transrectal ultrasound". *IEEE transactions on medical imaging*, 38(12), 2768-2778, 2019.
- [22] Zhao H, Zhang Y, Liu S, Shi J, Loy CC, Lin D, Jia J. "Psanet: Point-wise spatial attention network for scene parsing". In *Proceedings of the European Conference on Computer Vision (ECCV)*, Munich, Germany, 8-14 September, 2018.
- [23] Li H, Xiong P, An J, Wang L. "Pyramid attention network for semantic segmentation". arXiv, 2018. <https://arxiv.org/pdf/1805.10180.pdf>
- [24] Yu Y, Ji Z, Fu Y, Guo J, Pang Y, Zhang Z. "Stacked semantic-guided attention model for fine-grained zero-shot learning". *32<sup>nd</sup> Conference on Neural Information Processing Systems, Montreal*, Canada, 3-8 December, 2018.
- [25] Fu J, Liu J, Tian H, Li Y, Bao Y, Fang Z, Lu H. "Dual attention network for scene segmentation". *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Long Beach, CA 16-20 June, 2019.
- [26] Wu F, Chen F, Jing XY, Hu CH, Ge Q, Ji Y. "Dynamic attention network for semantic segmentation". *Neurocomputing*, 384, 182-191, 2020
- [27] Sinha A, Dolz J. "Multi-scale self-guided attention for medical image segmentation". *IEEE journal of biomedical and health informatics*, 25(1), 121-130, 2020
- [28] Huang Z, Wang X, Huang L, Huang C, Wei Y, Liu W. "Ccnet: Criss-cross attention for semantic segmentation". *2019 IEEE/CVF International Conference on Computer Vision*, Seoul, Korea, 27 October-2 November, 2019.

- [29] Tao A, Sapra K, Catanzaro B. "Hierarchical multi-scale attention for semantic segmentation". *arXiv*, 2020. <https://arxiv.org/pdf/2005.10821.pdf>
- [30] Chen LC, Yang Y, Wang J, Xu W, Yuille AL. "Attention to scale: Scale-aware semantic image segmentation". *2016 IEEE conference on computer vision and pattern recognition* Las Vegas, NV, USA, 27-30 June 2016.
- [31] Li R, Su J, Duan C, Zheng S. "Linear attention mechanism: An efficient attention for semantic segmentation". *arXiv*, 2020 <https://arxiv.org/ftp/arxiv/papers/2007/2007.14902.pdf>
- [32] Islam M, Vibashan VS, Jose V, Wijethilake N, Utkarsh U, Ren H. "Brain tumor segmentation and survival prediction using 3D attention UNet". *5<sup>th</sup> International Workshop, BrainLes 2019, Held in Conjunction with MICCAI 2019*, Shenzhen, China, 17 October 2019.
- [33] Zhang J, Yu L, Chen D, Pan W, Shi C, Niu Y, Cheng Y. "Dense GAN and multi-layer attention based lesion segmentation method for COVID-19 CT images". *Biomedical Signal Processing and Control*, 2021. <https://doi.org/10.1016/j.bspc.2021.102901>
- [34] Camelyon16 Veri Kümesi. <https://camelyon16.grand-challenge.org/> (08.05.2022)
- He K, Zhang X, Ren S, Sun J. "Deep residual learning for image recognition". *2016 IEEE conference on computer vision and pattern recognition*, Las Vegas, Nevada. USA, 27-30 June 2016.