## 基本資料:

標題是"MIDeepSeg: Minimally Interactive Segmentation of Unseen Objects from Medical Images Using Deep Learning"

一張含有 文字, 字型, 螢幕擷取畫面 的圖片

自動產生的描述

## 摘要:

醫學影像中器官或病變的分割在臨床應用中的重要性，以及傳統的捲積神經網路（CNN）在自動分割方面的限制。 雖然CNN在自動分割方面取得了最先進的效能，但在複雜情況下，它們通常受到臨床上可接受的準確性和穩健性的限制。 因此，互動式分割是這些方法的一個實際替代方案。 然而，傳統的互動式分割方法需要大量使用者交互，而最近提出的基於CNN的交互式分割方法在先前未見過的物件上表現較差。

為解決這些問題，論文提出了一種新穎的基於深度學習的互動式分割方法。 該方法具有高效性，因為它僅需要使用者點擊作為輸入，並且能夠在一系列先前未見過的物件上實現良好的泛化。 具體來說，該方法首先透過提出的指數化測地距離對使用者提供的內部邊緣點進行編碼，使CNN能夠獲得先前見過和未見過對象的良好初始分割結果，然後使用一種新穎的信息 融合方法，將初始分割與僅少量附加的使用者點擊相結合，以有效地獲得精細的分割。 作者透過在2D和3D醫學影像分割任務上進行了廣泛的實驗來驗證他們提出的框架，這些任務涵蓋了在訓練集中未出現的各種先前未見過的物件。 實驗結果表明，他們提出的框架在準確性方面比最先進的互動框架需要更少的使用者互動和更少的時間，並且對先前未見過的物件具有良好的泛化能力。

關鍵字: 互動式影像分割、卷積神經網路、測地距離、泛化

## 介紹

這篇論文的介紹部分討論了醫學影像中器官或病變的準確和穩健的分割在臨床應用中的重要性，尤其在診斷和治療規劃方面。 儘管自動分割方法已經研究了很多年，但在具有大量解剖變異和複雜病理情況的案例中，它們仍然難以獲得一致精確的分割結果。 這主要是由於醫學影像的固有限制，例如低對比度、不同的影像和分割協議以及患者之間的變異。 與此相反，互動式分割方法利用使用者的知識和經驗來獲得更準確和穩健的結果，因此在臨床應用中更為實際和廣泛使用。

理想的互動式分割工具應該具備以下特點：1）以盡量少的使用者輸入實現精確的分割結果，減輕使用者的負擔；2）具有高效性，以便使用者可以在處理體積資料時獲得即時回應； 3）對不同的物件具有良好的泛化能力，以便可以用於新的物件或影像模態。 然而，現有的互動式分割方法很少能夠同時滿足所有這些常常相互競爭的要求。 許多傳統的互動方法使用低階特徵（例如灰階或色彩分佈）進行影像分割，如圖割（Graph Cuts）、ITK-SNAP、GeoS、Random Walks和GrowCut等。 由於低階特徵在許多低對比情況下無法有效地區分目標和背景，因此這些方法通常需要大量的使用者互動和較長的使用者時間才能獲得可靠的結果。 為了減少使用者建立足夠前景/背景模型所需的註解數量，機器學習已廣泛用於執行互動式分割。 例如，SlicSeg和DyBaORF使用線上隨機森林（ORF）來分割磁振造影（MRI）體積中的胎盤。 GrabCut使用高斯混合模型（GMM）來估計前景和背景分佈。 它透過使用者提供的興趣區域周圍的邊界框獲得初始結果，並允許進一步的互動來進行細化。 Wang等人使用主動學習來主動選擇候選區域，以詢問用戶以獲得更多信息，從而減少用戶互動。 這些演算法相對於沒有機器學習的傳統方法表現更好，但它們仍然需要相當數量的用戶互動才能實現精確的分割。

最近，隨著深度學習和卷積神經網路（CNNs）能夠自動學習高級語義特徵，它們在影像分割方面取得了最先進的性能。 為了充分利用CNNs的良好表示能力並克服自動CNNs的有限準確性和穩健性，一些基於深度學習的互動式分割工具已經被提出。 然而，這些方法通常缺乏對具有低對比度和模糊邊界的醫學影像的評估。 總之，這個介紹部分提供了關於醫學影像分割的挑戰以及互動式分割方法的演變和當前的研究方向的綜述。

在相反的情況下，DeepIGeoS、IFSeg、DeepCut和BIFSeg等方法是專門設計用於醫學影像分割的。 DeepCut使用一組由使用者提供的邊界框作為稀疏註釋，用於訓練CNN來分割胎兒MRI中的胎兒大腦和肺部。 Roth等人將極端點與隨機行者結合，用於弱監督的3D醫學影像分割。 儘管這種方法和DeepCut顯著降低了註釋成本，但它是為了在大型資料集上進行弱監督模型訓練，而不是在測試時互動式地編輯單一分割結果而設計的。 Raju等人進一步使用極端點在病理性肝臟分割的使用者引導領域自適應方法中。 DeepIGeoS透過結合CNN和使用者提供的塗鴉實現使用者友善的互動式分割，其中一個CNN用於獲得初始分割，另一個CNN接受額外的使用者互動進行精細化。 然而，DeepIGeoS只能處理訓練集中存在的對象，並且缺乏對先前未見過的對象的適應性。 IFSeg採用使用者點擊和原始影像作為輸入進行醫學影像互動式分割，儘管該框架易於使用，但其泛化能力僅在單一先前未見過的結構上進行了驗證，未展示其處理不同模態中的 各種未見過對象的能力。 BIFSeg利用使用者提供的邊界框和影像特定的微調來分割一些未見過的對象，但它受限於處理同一影像模態或類似上下文中僅有少數未見過的對象，並需要耗時的微調 用於每個測試圖像。 因此，對於具有更有效率性和泛化能力的醫學影像分割的新型互動框架是非常值得的。

此外，基於CNN的交互式分割方法面臨的一個實際問題是如何有效地編碼使用者交互，因為不同的編碼策略對互動式分割效能有很大影響。 大多數現有方法透過將使用者互動轉換為提示圖（cue map）來編碼它們，例如歐幾里德距離圖、高斯熱度圖和從使用者點擊中派生的等值輪廓。 然而，這些編碼方法不考慮圖像上下文資訊。 相反，測地距離變換對於編碼使用者互動是空間平滑且對比敏感的（Criminisi等，2008; Bai和Sapiro，2009; Price等，2010）。 DeepIGeoS使用測地距離變換以特殊設計的閾值處理使用者提供的互動。 然而，在處理不同物件時，找到適當的截斷產生的測地距離圖的閾值值是耗時的。 作者認為，一種具有上下文感知性和無需參數的編碼方法有助於提高分割的準確性和泛化能力。

為了應對上述挑戰，我們提出了一個新的通用框架，用於更聰明和準確地互動分割2D和3D醫學影像。 該框架旨在不僅獲得先前見過物件的高效能和高效性，還在一系列先前未見過的物件上實現高泛化能力。 我們的方法充分利用了CNN，並且僅需要少量使用者點擊作為使用者互動。 我們提出了一種基於指數化測地距離（EGD）變換的新方式來編碼使用者交互，這種方式具有上下文感知性和無需參數，有助於改善CNN獲得的分割結果。 我們還提出了一種資訊融合方法，將額外的使用者點擊與初始分割高效融合，以獲得精細的分割結果。 與現有的互動式醫學影像分割框架不同，我們的方法更加高效，因為它僅在影像的子區域上運行，並且不需要即時訓練額外的CNN來進行細化。 此外，我們透過大範圍的先前見過和未見過的物件驗證了該框架的有效性。 我們的方法在不同類型的影像背景和模態中，透過驗證了2D未見過物件的五種類型和3D未見過物件的四種類型，證實了其優越性，相對於現有的互動式 分割方法。

圖1展示了提出的"Minimally Interactive Deep learning-based Segmentation framework"（MIDeepSeg）的流程圖。 這個框架包括兩個階段：

階段1：使用者提供的內部邊緣點透過指數化測地距離（EGD）映射進行編碼，以引導CNN獲得初始分割結果。

階段2：在額外的使用者點擊和作者提出的資訊融合後，採用圖割（Graph Cut）來對初始分割進行精細化。 需要注意的是，這個框架可以用於對先前未見過的物件進行分割，無需額外的微調或重新訓練。

圖2展示了在不同胎盤形狀的訓練影像上模擬的內部邊緣點。 其中，紫紅色表示在胎盤邊緣上模擬的點擊點，棕色表示從內部邊緣點派生的放鬆的邊界框，黃色表示地面真實標籤。

這些圖示幫助說明了MIDeepSeg框架的工作原理以及內部邊緣點如何用於引導分割和精細化過程。

## 方法:

在這個方法部分，描述了所提出的"Minimally Interactive Deep learning-based Segmentation framework"（MIDeepSeg）的兩個階段以及使用者互動的基礎。

首先，MIDeepSeg框架包括兩個階段。 在第一階段，使用者需要在目標物件的邊界附近提供少量點擊（即內部邊緣點）。 這些點用於推斷一個放鬆的邊界框，以裁切輸入影像。 基於裁切後的影像，所有使用者提供的內部邊緣點都會轉換為基於作者提出的EGD變換的提示圖。 然後，提示圖與裁剪後的輸入影像連接，作為CNN的輸入，以獲得初始分割結果。 在第二階段，使用者提供額外的點擊，用於指示錯誤分割的區域，並透過作者提出的資訊融合和圖割（IF-GC）來獲得精細化的結果。 在測試時，精細化步驟可以運行多次，直到使用者接受結果。 在用一組物件進行訓練後，我們的框架可以用於對先前未見過的物件進行分割，而無需耗時的微調或重新訓練以獲取額外的註釋。

接下來，介紹了基於內部邊緣點的使用者互動。 許多現有的基於CNN的互動式分割框架使用塗鴉或邊界框作為互動提示，這需要使用者小心拖曳遊標，需要使用者大量的努力。 與之不同，使用點擊作為用戶互動方式更加用戶友好和有效，正如以前的研究所證明的那樣。 然而，在醫學影像中，精確的極端點很難找到，而且對不同患者或影像協議的不同器官或病灶在大小和形狀上存在較大的變化，特別是在3D體積資料中。 此外，對於不規則和凹凸不平的形狀，僅使用極端點無法捕捉物體的主要形狀，這可能會限制CNN的性能。

為了克服這些限制，我們建議使用內部邊緣點作為使用者互動的方式，使用者只需提供位於目標邊界內部並且靠近邊界的一些點擊。 與只使用最多四個極端點的DEXTR方法不同，我們的內部邊緣點可以為具有複雜和不規則形狀的不同類型器官提供更多形狀資訊。 此外，將點擊放在物件邊界上甚至極端點上對用戶來說在測試時很困難，而將點擊放鬆到邊界內部使得互動更加友好和方便實施，容忍不準確的點擊。 因為內部邊緣點的指數化測地距離變換可以很好地近似目標物件的顯著性圖，所以內部邊緣點在引導CNN處理不同類型未見過物件方面具有潛在的優勢。

在訓練過程中，我們自動模擬了每個物件的所有內部邊緣點，這是基於地面真實掩膜和邊緣偵測器（Harris等，1988）進行的。 內部邊緣點的產生基於兩個規則：首先，這些點應位於物件內部並靠近邊界。 其次，由這些點決定的放鬆邊界框應覆蓋整個物件區域。 因此，我們在訓練圖像上以兩個步驟模擬使用者互動：

1) 為確保放鬆邊界框覆蓋感興趣區域的整個區域，選擇地面真實邊界上靠近目標對象的極端點附近的少量點（對於2D對象選擇三或四個點，對於3D對象選擇五或六個點 ）。 然後，從目標的剩餘邊界點中隨機採樣n個點，以提供更多的形狀信息，其中n是從0到5的隨機數字。

2) 為了模擬可能不準確地定位在物件邊界上的真實使用者點擊，將在步驟1中獲得的所有這些點向邊界內側移動幾個像素/體素，以獲得我們的內部邊緣點。 我們將模擬的點朝向目標物件的內部移動，因為使用者也被要求將內部邊緣點放在邊界的內側。 然後，由這些點決定的邊界框會被放鬆，以包含一些背景區域。 圖2顯示了在訓練影像上模擬的2D內部邊緣點和放鬆的邊界框範例。

在測試階段，使用者需要以滿足上述兩個規則的方式提供內部邊緣點。 由使用者互動決定的放鬆邊界框會在其周圍擴展一小部分，以包含一些上下文資訊。

另外，對於編碼使用者互動的效率至關重要。 理想的編碼方法應考慮影像上下文，並可直接與CNN結合，無需手動設計參數。 然而，現有的交互編碼方法，如歐幾里德距離變換、高斯熱度圖、等值輪廓和測地距離變換，不同時具備這些優點。 因此，為了解決這個問題，我們提出了一種具有上下文感知性和無需參數的編碼方法：指數化測地距離（EGD）變換，它是測地距離變換和指數變換的組合。

在這個部分中，描述了用於計算指數化測地距離（EGD）的數學公式和相關概念。

假設Ss代表訓練階段中模擬的內部邊緣點或測試階段中使用者提供的內部邊緣點所屬的像素/體素集合。 對於輸入影像I中的一個像素/體素i，從i到Ss的無符號EGD定義如下：

EGD(i, Ss, I) = min

j∈Ss

e

−Dgeo(i,j,I)

(1)

其中Dgeo(i, j, I)是從像素/體素i到像素/體素j的測地距離，定義如下：

Dgeo(i, j, I) = min

p∈Pi,j Z 1

0

k∇I(p(n)) · v(n)k dn (2)

這裡，Pi,j是像素/體素i和j之間所有路徑的集合，p是一條可行路徑，由n ∈ [0, 1]參數化。 v(n) = p

0

(n)/ kp

0

(n)k 是沿路徑方向的單位向量。 要注意的是，這裡定義的EGD適用於標量影像，但可以輕鬆擴展到向量值影像（即多通道或多模態影像）。 圖3顯示了應用於一些內部邊緣點的不同編碼方法所得到的提示圖的範例。 可以觀察到，基於EGD的提示圖能更好地區分前景和背景，比其他編碼方法更好。 因此，它具有潛力提供更多的形狀、位置和上下文信息，以引導CNN獲得良好的初始分割結果。

在這篇論文中，我們專注於設計一個高效和通用的框架，以處理來自不同類型圖像的已知和未知物件。 因此，我們的框架不依賴特定的CNN結構設計。 為了展現其實用性，我們分別使用了適應的2D-UNet（Ronneberger等，2015）和3D-UNet（C¸i¸cek等，2016）進行2D和3D分割。 我們將批量歸一化層替換為實例歸一化層，這對不同類型影像具有更好的適應性，並將特徵通道數減少了四倍，以平衡效能、記憶體成本和時間消耗。 在訓練階段，根據地面真實標籤，自動模擬了所有內部邊緣點和放鬆的邊界框，如第2.1節所述。 然後，所有內部邊緣點都轉換為提示圖，與裁剪後的輸入影像連接在一起，作為CNN的輸入，如圖2所示。 在測試階段，使用者需要為給定的目標提供內部邊緣點。 然後，CNN可以給出一個初始分割結果。 為了修正誤分割，我們使用了一個包含初始分割和額外使用者點擊資訊融合的細化階段，如下所述。

2.4. 基於初始分割和額外使用者點擊的資訊融合的細化

對於基於深度學習的互動式分割，支援初始分割的細化是非常重要的。 現有方法要麼需要一個額外的模型進行細化（Wang等人，2018b；Castrejon等人，2017；Acuna等人，2018；Zhou等人，2019a；Liao等人，2020），要麼需要對特定影像進行 微調（Wang等人，2018a）。 然而，這些細化方法既耗時又佔用內存，並且不適用於未知物件。 此外，Chen等人（2017）和Kamnitsas等人（2016）使用了CRF（Lafferty等人，2001）來自動細化CNN的預測。 然而，這些基於CRF的（Lafferty等人，2001）細化方法（Chen等人，2017；Kamnitsas等人，2016）並不是為互動式分割而設計的。 與這些方法不同，我們提出了一種高效和簡單的細化方法，基於初始分割和額外用戶互動之間的資訊融合的新方法，它在不需要額外微調和重新訓練的情況下更好地推廣 到以前未見的物件。 圖4顯示了我們資訊融合方法的示意圖。

在細化階段，使用者需要提供額外的點擊來指示被錯誤分割的前景和背景區域。 為了有效率地編碼這些新的交互，我們再次使用了提出的EGD變換，以獲得兩個額外的交互導出的提示圖：Ef和Eb是用於細化的基於用戶提供的前景和背景點擊的EGD的 提示圖。 請注意，我們不會直接重複使用第一階段獲得的初始EGD圖，而是將初始的內部邊緣點與細化點擊組合起來，用於在細化步驟中計算新的EGD圖。 Ef和Eb的值在[0, 1]範圍內，表示每個像素與前景/背景點擊之間的相似性。 設Pf和Pb分別表示CNN所得的初始前景和背景機率圖，i表示輸入影像I中的像素/體素。 資訊融合策略旨在根據Ef和Eb來細化Pf和Pb。 具體來說，我們的目標是在像素i接近細化點擊時自動強調Ef和Eb，否則Pf和Pb趨向於保持不變。 我們定義了像素i的使用者校準前景（Rfi）和背景（Rbi）機率如下：

Ef

i

=

e

−Dfi

e

−Dfi

+

e

−Dbi

（3）

Eb

i

=

e

−Dbi

e

−Dfi

+

e

−Dbi

（4）

Rfi

=

（1 − αi）∗Pfi

+

αi∗Ef

i

（5）

Rbi

=

（1 − αi）∗Pbi

+

αi∗Eb

i

（6）

αi

=

e

− min（Dfi，Dbi）

（7）

其中αi∈[0，1]是自動的自適應權重因子。 當i接近點擊時，αi接近1.0，Rfi（Rbi）更受Ef

i

（Eb

i

）的影響。 如果沒有為前景（背景）提供點擊，我們將相應的Dfi

i

或Dbi

i

設定為常數值。 讓C

f和C

b分別表示前景和背景的點擊，所以點擊的整個集合是C = C

f∪C

b。 讓ci表示點擊中像素的使用者提供的標籤，然後我們有ci = 1如果i ∈ C

f，並且ci = 0如果i ∈ C

b。 我們將Rfi和Rbi整合到條件隨機場（CRF）中，以獲得細化分割：

E =

P

i

φ(yi|I) + λ∗

P

i,j

ψ(yi

圖4：資訊融合的細化示意圖。 (a) 使用者提供點擊來指示欠分割（紅色）和過分割（青色）區域。 (b) 和 (c) 是在第一階段由CNN所獲得的初始分割前景和背景機率圖，分別。 (d) 和 (e) 是基於前景和背景細化點擊和EGD變換的提示圖，分別。 (g) 和 (h) 是校準的前景和背景機率圖，分別。 (f) 是細化的分割結果。 (IF: 資訊融合; EGD: 指數化的測地距離轉換; GC: 圖割)

## 實驗與結果

3. 實驗與結果

3.1. 比較方法與評估指標

為了研究不同編碼方法在我們分割方法的第一階段中使用相同的內部邊緣點時的性能，我們將我們的EGD與歐幾 里德距離變換、高斯距離變換和測地距離變換進行了比較，分別 稱為EGD、Eucl、Gauss和Geos。 此外，我們也將它們與不編碼交互作用的邊界框分割方法進行了比較，稱為BBox。 所有這些方法都基於相同的CNN結構。 為了公平比較，Eucl、Gauss和Geos是根據其用於編碼使用者提供的互動的最佳參數進行實現的。 （請參閱補充文件）

MIDeepSeg也與幾種現有的互動式分割方法進行了比較。 在2D情況下，除了傳統方法如圖割（Boykov和Jolly，2001）、隨機遊走（Grady，2006）和SlicSeg（Wang等，2016b）之外，我們還比較了最近的基於深度學習的方法， 包括DeepIGeoS（Wang等，2018b）、DIOS（Xu等，2016）、DeepGrabCut（Xu等，2017）和DEXTR（Maninis等，2018），其中使用了與我們2D版本MIDeepSeg相同的2D網路結構。 對於3D分割，我們將MIDeepSeg與ITK-SNAP（Yushkevich等，2006）和3D圖割（Boykov和Jolly，2001）以及DeepIGeoS（Wang等，2018b）、DIOS（Xu等，2016）、Deep等GrabCut（Xu，等GrabCut（Xu等，2016）、Deep等GrabCut（Xu，等GrabCut（Xu等，2016）、Deep等GrabCut（Xu等， 2017）和DEXTR（Maninis等，2018）的3D版本進行了比較，這些方法使用了與MIDeepSeg相同的3D網路進行3D分割。 圖割、SlicSeg、隨機遊走、DeepIGeoS和DIOS讓使用者多次細化結果。 DeepGrabCut只允許使用者在開始時繪製邊界框，並不支援進一步的細化互動。 DEXTR將極端點作為使用者交互，並允許使用者細化結果一次。 圖割、SlicSeg、隨機遊走和ITK-SNAP是傳統的互動式分割方法，無需使用標註資料進行訓練，具有很高的泛化性。 相反，DeepIGeoS、DIOS、DeepGrabCut和DEXTR是基於深度學習的方法，需要標籤的資料進行訓練，而DeepIGeoS無法處理未見過的物件。 兩位使用者分別使用這些互動框架來分割每個測試影像，直到結果在視覺上可接受，我們報告了兩位使用者達到的平均結果。 分割結果與由經驗豐富的放射科醫生手動註釋的地面真實標籤進行比較。 為了進行定量評估，我們使用了Dice相似係數和平均對稱表面距離（ASSD）。

Dice =

2· | Rp ∩ Rg |

| Rp | + | Rg |

(12)

其中，Rp和Rg分別表示由演算法分割的區域和地面真實標籤。

ASSD =

1

| Sp | + | Sg |

X

i∈Sp

d(i, Sg) + X

i∈Sg

d(i, Sp)

(13)

其中，Sp和Sg分別表示由演算法提供的結果的表面點集和地面真實標籤，d(i, Sp)表示點i與表面Sg之間的最短歐幾里德距離。 為了研究這些方法的效率，我們列出了每個分割任務的使用者時間和使用者互動點的數量。

物體 模態 訓練 測試 資料集

胎盤 MRI（T2） 532張切片 176張切片 我們的

脾臟 CT 235張切片 159張切片 BTCV，TCIA3

腎臟 MRI（T1） 無 100張切片 CHAOS

腎臟 MRI（T2） 無 100張切片 CHAOS

腎臟 CT 無 100張切片 CHAOS

脾臟 MRI（T1） 無 100張切片 CHAOS

脾臟 MRI（T2） 無 100張切片 CHAOS

脾臟 CT 無 100張切片 CHAOS

攝護腺 MRI（T2） 無 72張切片 MSD

胎兒腦部超音波 無 60張切片 HC18

表1：用於訓練和測試2D互動式分割框架的資料集。 請注意，對於脾臟，BTCV和TCIA分別是訓練集和測試集。

在第3.2節中，我們首先使用兩個2D應用來驗證提出的流程：從胎兒MRI中分割胎盤和脾臟的分割，分別來自腹部CT。具體來說，胎盤數據來自於第二季度的30個懷孕的臨床MRI掃描，以軸向視圖獲取，像素大小介於0.7422 mm × 0.7422 mm和1.582 mm × 1.582 mm之間，切片厚度為3 - 4 mm。每個切片都被重新採樣為統一的像素大小1 mm×1 mm。我們使用了來自18個體積的532個切片進行訓練，來自4個體積的111個切片進行驗證，來自8個體積的176個切片進行測試。地面真實值由經驗豐富的放射科醫師手動劃定。對於脾臟數據，我們從BTCV（Marsh, 2013）數據集中選擇了47個體積（每個體積5個切片）的235個脾臟切片進行訓練，並從TCIA4數據集中選擇了53個體積（每個體積3個切片）的159個脾臟切片進行測試。其次，為了驗證我們方法的通用性，我們將僅在MRI中的胎盤訓練的模型應用於訓練集中未包含的多種模式中出現的四種器官：1）CT中的腎臟，CHAOS5訓練集中的T1加權和T2加權MRI。我們分別隨機選擇了這三種情況的100個切片。2）CHAOS訓練集中的CT、T1加權和T2加權MRI中的脾臟。我們也分別隨機選擇了這三種情況的100個切片。3）來自MSD6 Task05數據集的T2加權MRI中的前列腺，我們隨機選擇了24個案例中的72個切片。4）來自HC187數據集的超聲影像中的胎兒大腦，我們隨機選擇了60個切片。訓練集和測試集的信息在表1中列出。為了處理不同尺度的不同器官，我們將裁剪的子區域和提示圖調整為64×64，作為CNN的輸入。

在第3.2.2節中，我們展示了使用提供的內部邊緣點的胎盤（來自MRI）和脾臟（來自CT）的初始分割的一些示例。我們分別將提出的EGD與BBox、Eucl（帶有閾值）、Gauss（帶有sigma）、Geos（帶有閾值）進行了比較，並使用相同的用戶提供的內部邊緣點。請注意，Eucl、Gauss和Geos的參數分別經過了優化以進行比較，更多的詳細信息列在附錄的第8節中。可以觀察到，EGD變換可以引導CNN獲得比其他編碼方法更準確的分割結果。表2列出了不同編碼方法對胎盤和脾臟的定量評估結果。可以觀察到，我們的上下文感知和無參數的EGD編碼方法一致優於其他方法。2D中的EGD的計算時間小於0.05秒，可實時響應。圖6展示了對具有複雜形狀的挑戰性案例的初始分割的不同內部邊緣點數量的影響。

在第3.2.3節中，我們展示了使用不同的精緻方法進行胎盤和脾臟分割的交互式精緻示例。第一行顯示了在我們框架的第一階段獲得的初始分割。基於初始分割，我們進一步使用附加的點來獲得精緻的結果。我們比較了使用相同用戶點的不同精緻方法之間的精緻結果，包括純粹的圖割（GC）和信息融合後的圖割（IF-GC）。根據BIFSeg（Wang et al., 2018a）中的實現，純粹的圖割將初始分割概率圖和用戶交互（背景和前景種子）作為輸入，通過最大流解決。在表3中列出了胎盤和脾臟分割的性能，前兩行顯示我們方法在第一階段已經遠遠優於具有相同網絡結構的自動分割。最後兩行演示了我們的IF-GC在第二階段的精緻中比純粹的圖割在相同的用戶點的精緻方面實現了更高的準確性。我們進一步研究了使用MIDeepSeg進行不同對象的分割所需的精緻點的數量，並繪製了精緻點數量的直方圖，如圖8所示。可以發現，許多測試案例不需要額外的點來獲得準確的結果，只有一些具有挑戰性的案例需要超過4個點進行精緻。

3.2.4

我們將MIDeepSeg與DeepIGeoS（Wang等人，2018b），圖割（Boykov和Jolly，2001），隨機遊走（Grady，2006），SlicSeg（Wang等人，2016b），DIOS（Xu等人，2016）， DeepGrabCut（Xu等人，2017）和DEXTR（Maninis等人，2018）進行了比較，分別用於胎盤和脾臟的分割。 圖9顯示了這些方法在2D胎盤分割上的可視比較。 第一行顯示了初始互動和初始分割結果，第二行顯示了精化後的最終結果和所有使用者互動。 可以看出，MIDeepSeg只需較少的使用者點擊就可以獲得良好的結果，而其他方法需要更多的互動。 根據胎盤和脾臟的結果，表4中的這些方法的定量比較顯示了MIDeepSeg的精度更高，與其他方法相比，它需要更少的用戶時間和較少的交互點，除了DeepGrabCut。 請注意，DeepGrabCut不允許進一步的使用者互動以進行精化，這導致了與其他方法相比最低的精確度。 這表明我們的方法非常高效，可以獲得高度準確的分割結果。

3.2.5

為了調查MIDeepSeg在以前未見過的物件上的表現和通用性，我們將MIDeepSeg與具有良好通用性的現有方法進行了比較：圖割（Boykov和Jolly，2001），隨機遊走（Grady，2006 ），SlicSeg（Wang等人，2016b），DIOS（Xu等人，2016），DeepGrabCut（Xu等人，2017）和DEXTR（Maninis等人，2018）。 對於MIDeepSeg，DIOS，DeepGrabCut和DEXTR，我們使用了僅使用胎盤圖像（T2加權MRI）訓練的模型來分割四個以前未見過的器官（即腎臟、脾臟、前列腺和胎兒大腦）以各種模態 ，如表1所列。 圖10顯示了MIDeepSeg對先前未見過的物件的分割結果範例。 第一行顯示了初始互動和初始分割結果。 在第二行中，呈現了所有交互作用和最終分割結果。 可以看出，MIDeepSeg只需很少的用戶點擊即可在未見過的器官上獲得良好的結果。 基於最終結果的這些方法的定量比較如圖11所示。 它顯示MIDeepSeg需要明顯較少的使用者時間和交互，但與其他互動式分割方法相比，具有相似或更高的準確性。 此外，可以觀察到，MIDeepSeg可以很好地處理不同類型的以前未見過的圖像模態和器官，而無需任何額外的訓練或微調。 我們進一步研究了使用MIDeepSeg進行腎臟（T2-MRI）和脾臟（T2-MRI）分割的精化點擊數量，並繪製了精化點擊數量的直方圖，如圖8所示。 我們可以發現，儘管這些物件不在訓練集中，但我們的方法需要很少或僅需少量點擊即可精細調整以獲得準確的結果。

3.3

3.3.1

首先，我們驗證了MIDeepSeg在從增強對比T1加權影像中分割3D腦腫瘤核的表現。 我們使用了BraTS20189訓練集，該訓練集包含285個帶有四種模態的案例：FLAIR、T1ce、T1和T2。 所有影像都已移除顱骨並重新取樣為1mm × 1mm × 1mm的等軸解析度。 我們使用了170個T1ce案例進行訓練和47個T1ce案例進行測試。 手動分割用作地面實況。

然後，我們驗證了MIDeepSeg的通用性，透過三個未見過對象的分割任務：1）來自BraTS2018的FLAIR中的整個腦腫瘤，我們隨機選擇了60個案例進行測試。 2）來自KiTS201910資料集的CT中的腎臟，在這裡我們隨機選擇了15個案例（包括30個有或沒有腫瘤的腎臟）進行測試。 3）來自ACDC11的MRI中的左心室，我們隨機選擇了30個案例。 KiTS和ACDC的測試數據已重新採樣為1mm × 1mm × 1mm的等軸分辨率。 所有用於訓練和測試的資料集都列在表5中。 為了處理不同尺度的3D對象，我們將裁剪的子區域和提示地圖調整為64 × 96 × 96。

3.3.2

為了驗證我們提出的EGD轉換在3D體積中的內部邊緣點編碼中的有效性，我們將其與BBox、Eucl、Gauss和Geos進行了比較，分別使用了它們各自優化的參數。 在這個階段，使用了使用者提供的相同內部邊緣點集合來進行這些方法的比較。 圖12顯示了不同編碼方法引導的CNN所獲得的初始分割結果。 可以看到，EGD轉換可以引導CNN從BBox中獲得更顯著的改進，與其他編碼方法相比。 表6列出了從T1ce影像中分割腫瘤核的不同編碼方法的定量評估結果。 可以看到，我們的上下文感知和無參數編碼方法EGD在腫瘤核的Dice和ASSD方面都表現出色，分別為87.00%和1.46 mm。 儘管EGD需要更多的時間（0.24秒）來進行互動編碼，但在實踐中仍然非常有效率。

3.3.3

基於上述由我們的方法獲得的初始分割結果，我們進一步使用額外的點擊來獲得精細化的結果。 我們將使用相同的使用者精化點擊集合來比較天真的圖割（GC）和提出的資訊融合後的圖割（IF-GC）之間的精化結果。 腫瘤核分割的性能列在表7中，顯示資訊融合比其他變異體獲得了更高的準確性。

3.3.4

圖14顯示了MIDeepSeg、3D圖割（Boykov和Jolly，2001）、ITKSNAP（Yushkevich等人，2006）以及DeepIGeoS（Wang等人，2018b）、DIOS（Xu等人，2016）、DeepGrabCut（Xu等人， 2017）和DEXTR（Maninis等人，2018）的視覺比較。 可以看出，MIDeepSeg只需要少量內部邊緣點作為初始交互，但其初始分割更準確，需要更少的使用者點擊才能獲得準確的最終結果。 基於最終結果的這些方法的定量比較見表8。 它顯示MIDeepSeg的準確性明顯高於其他方法。 此外，MIDeepSeg平均需要29秒來完成腫瘤核的整個3D互動分割過程，比DeepGrabCut以外的其他方法都要少。 （見附錄影片第C部分）

3.3.5

為了探討MIDeepSeg在先前未見過的3D物件上的泛化能力，我們使用了從T1ce影像中的腫瘤核分割任務中訓練的3D CNN模型，處理了三種以前未見過的物件和模態性 ：FLAIR中的整個腫瘤；CT中的腎臟和MRI中的左心室，如表5所示。 兩位使用者使用了MIDeepSeg和兩種具有良好泛化性的現有方法，包括ITK-SNAP（Yushkevich等人，2006）和Graph Cuts（Boykov和Jolly，2001）的3D版本，DIOS（Xu等人， 2016），DeepGrabCut（Xu等人，2017）和DEXTR（Maninis等人，2018），用來分割這些物體。 圖15顯示了使用MIDeepSeg分割的一些3D整個腫瘤、腎臟和左心室的範例。 可以看到，使用MIDeepSeg並進行少量點擊即可獲得不同類型的未見過物件的準確結果。 定量評估結果如圖16所示。 它顯示MIDeepSeg在與3D圖割、ITK-SNAP、DeepIGeoS、DIOS、DeepGrabCut和DEXTR相比實現了類似或更高的準確性。 然而，MIDeepSeg需要更少的使用者時間來實現這些結果。 （見附錄影片第D部分）

圖表

為了研究MIDeepSeg對以前未見過的3D對象的泛化能力，我們使用了從T1ce圖像中的腫瘤核分割任務中訓練的3D CNN模型，處理了三個以前未見過的對象和模式：FLAIR 中的整個腫瘤；CT中的腎臟和MRI中的左心室，如表5所列。 兩位用戶使用了MIDeepSeg和兩種具有良好泛化能力的現有方法，包括ITK-SNAP（Yushkevich等，2006）和Graph Cuts（Boykov和Jolly，2001）的3D版本，DIOS（Xu等，2016）， DeepGrabCut（Xu等，2017）和DEXTR（Maninis等，2018）來分割這些物件。 圖15顯示了使用MIDeepSeg進行的3D整個腫瘤、腎臟和左心室分割的一些例子。 可以看到，使用MIDeepSeg僅需少量點擊即可獲得不同類型的未見過對象的準確結果。 定量評估結果如圖16所示。 它顯示MIDeepSeg在與3D圖割、ITK-SNAP、DeepIGeoS、DIOS、DeepGrabCut和DEXTR相比實現了類似或更高的準確性。 然而，MIDeepSeg需要較少的使用者時間來實現這些結果。 （見附加影片的D部分）

以下是您提供的每個圖表和表格的簡要描述：

圖5：我們方法的第一階段，用於胎盤和脾臟分割的不同編碼方法的視覺比較。第一列顯示了帶有用戶提供的內部邊緣點（品紅色）的輸入圖像。其他列顯示了初始結果。

圖6：不同數量的初始內部邊緣點對於胎盤（已知對象）和前列腺（未知對象）分割的影響，前列腺具有複雜的形狀。第一行顯示了具有不同數量內部邊緣點的輸入圖像。第二行顯示了分割結果。

表2：使用相同內部邊緣點集的不同編碼方法對胎盤和脾臟分割的定量比較。這些結果基於我們框架的初始分割。∗表示與第二名方法比較時的p值<0.05。

圖7：GC和IF-GC的視覺比較。第一行顯示了用於精緻初始分割結果的用戶點擊。其他行顯示了GC和IF-GC的精緻結果，這些結果基於相同的用戶點擊進行精緻。 （GC：純粹的圖割，IF-GC：信息融合後的圖割）

圖8：MIDeepSeg需要的不同對象的精緻點數量直方圖。第一行中的胎盤（MRI）和脾臟（CT）是已知對象，而第二行中的脾臟（T2-MRI）和腎臟（T2-MRI）是以前未見過的對象。

表3：使用相同的點擊集對胎盤和脾臟分割的不同精緻方法的定量比較。GC：純粹的圖割；IF-GC：信息融合後的圖割。\*表示與GC相比具有顯著差異（p值<0.05）。

表4：不同交互式方法的2D胎盤和脾臟分割的定量比較，包括Dice、ASSD、用戶時間和交互點數。∗表示與第二名方法比較時的p值<0.05。

\*\*圖9：\*\* 這張圖比較了MIDeepSeg和其他2D胎盤分割的交互方法。 第一行顯示了具有或不具有初始交互作用的初始分割結果。 第二行顯示了精煉後的最終結果。

\*\*圖10：\*\* 這些範例展示了MIDeepSeg進行的2D未見過的器官分割結果。 第一行顯示了初始用戶互動和初始分割。 第二行顯示了所有使用者互動和最終分割結果。 請注意，該模型僅在T2 MRI中訓練有胎盤。

\*\*表5：\*\* 這個表格列出了用於3D實驗的訓練和測試的數據集。 Ntrain和Ntest分別表示用於訓練和測試的捲數。

\*\*圖11：\*\* 這張圖顯示了不同交互分割方法對未見過的對象的Dice、ASSD、用戶時間和交互點的影響。 其中，CT影像、超音波影像、T1加權MR影像和T2加權MR影像分別用於表示。 所有這些器官在訓練集中都是以前從未見過的。

\*\*圖12：\*\* 這張圖比較了不同編碼方法對3D腫瘤核分割的影響，該分割基於第一階段獲得的初始分割。 所有這些方法都使用相同的內部邊緣點和推斷的邊界框作為輸入圖像。

\*\*圖13：\*\* 這是一個使用不同精煉方法進行腫瘤核分割的範例。 可以觀察到在相同的精煉點的情況下，IF-GC比GC更準確地精煉了結果。

\*\*表6：\*\* 這個表格比較了不同編碼方法對3D腫瘤核分割的影響，這些方法都使用相同的內部邊緣點。 結果是基於我們框架的初始分割（第一階段）得出的。 \*表示與第二名方法相比有顯著的差異（p值<0.05）。

\*\*圖14：\*\* 這張圖比較了使用MIDeepSeg、DeepIGeoS、3D圖割和ITK-SNAP進行的3D腫瘤核分割。

\*\*表8：\*\* 這個表格提供了不同交互方法在3D腫瘤核分割方面的定量評估，包括Dice、ASSD和用戶時間。 \*表示與第二名方法相比有顯著的差異（p值<0.05）。

如果您對其中任何一個圖表或表格需要更多詳細信息，請隨時提出。

討論

討論部分討論了您提到的論文中的主要發現和貢獻。 以下是討論部分的主要內容：

- \*\*對現有工作的挑戰：\*\* 討論部分開始指出了當前深度學習交互分割方法存在的一個主要挑戰，即難以在以前未見過的對像類別上實現良好的泛化。 這是因為這些方法需要大量的已標記圖像來訓練，而醫學圖像的精確標註非常珍貴且稀缺。 這導致了CNN在處理訓練集中未出現的未見對象時性能受限。

- \*\*MIDeepSeg的優勢：\*\* 討論進一步介紹了MIDeepSeg方法的主要優勢。 與傳統的CNN和遷移學習方法不同，MIDeepSeg的一個主要優點是它能夠在不重新訓練或微調的情況下分割未見過的物件。 這降低了收集和標註數據的負擔，並且可以直接應用於分割或標註未見過的對象。

- \*\*MIDeepSeg的實現細節：\*\* 討論部分解釋了MIDeepSeg中提出的內部邊緣點（EGD）方法的實現細節。 EGD與地幾距離方法不同，它是無參數的，具有更高的泛化能力，並且自然地輸出機率圖，可用作用戶互動所指示的前景或背景的機率。

- \*\*計算效率：\*\* 討論部分提供了MIDeepSeg中各個階段的計算時間信息，包括EGD的計算時間，2D和3D CNN的推斷時間，以及CRF優化時間。 結果表明，MIDeepSeg在用戶互動的情況下具有良好的計算效率，適合實時交互分割未見過的對象。

總之，討論部分強調了MIDeepSeg方法的創新性和實用性，特別是在處理未見過的物件時，無需重新訓練或微調模型。 該方法的計算效率也使其成為一個有潛力的工具，用於醫學圖像的交互分割。

在實驗中，我們發現基於校準的機率圖和Graph Cuts的細化方法在不同器官和模態下的各種情況下均表現良好。 其優點包括：

1）細化步驟與基於CNN的初始分割步驟分離，因此可作為一個通用的細化工具，用於交互地修正不同網絡獲得的分割結果以及未見過的對象。

2）計算效率高，允許即時響應用戶交互，這對於提高交互分割的用戶體驗非常重要。

3）用戶互動被用作硬約束，確保用戶給定的點在細化後會具有所需的標籤。

然而，可能存在一個潛在問題，即在復雜情況下，需要相對較多的點擊才能獲得準確的結果。 然而，在實際應用中，我們的方法易於使用並且在處理不同未見過的對象時非常高效，如實驗結果所示。

一個一般性的問題是，互動式分割的結果可能取決於使用者的知識和經驗，因為使用者可以根據其主觀感覺不斷細化分割，直到其在視覺上可接受。 然而，我們的方法對用戶互動有一些要求：在第一階段，交互需要在內側邊界附近給出，而在第二階段，交互僅在不正確的區域給出，對於大多數情況來說，不 正確的區域很小，這限制了不同用戶提供的點擊範圍。 因此，我們的方法之間的用戶變化很小。 由於我們的方法不要求用戶精確地在邊界或極端點上點擊，我們的內部邊緣點容忍不準確的點擊，這更加用戶友好。 如圖10所示，在第一列中，內部邊緣點不精確且遠離邊界，在第四列中，頂部點也不準確，甚至被點擊到背景中，但它們仍然產生了良好的初始分割結果 。 這進一步證明了MIDeepSeg的魯棒性和泛化性。

最近，一些研究使用Fisher資訊、自然語言、主動學習和深度強化學習來開發智慧互動分割或標註工具。 在未來，使用主動學習和深度強化學習以及不確定性估計來引導用戶進行細化的互動方法可能具有潛在的改進互動分割效率的潛力。

## 總結

總結來說，本文提出了一個基於深度學習的互動式框架，對於醫學影像分割具有良好的泛化性，並且只需要用戶提供少量點擊作為輸入。 我們提出了一種新的上下文感知和無需參數的編碼方法，用於編碼用戶交互，以引導CNN進行良好的初始分割。 基於編碼方法，我們還提出了一種有效的細化方式，以提高分割結果的準確性。 該框架旨在提高對未見過對象的泛化能力，這對於基於深度學習的模型非常重要。 在各種2D和3D影像中對先前看到和未見過的多種器官或病變進行分割的實驗表明：

1）我們的內部邊緣點和基於EGD變換的框架在準確性和效率方面優於現有的基於深度學習的互動式分割工具。

2）所提出的框架在先前未見過的對像上具有很好的泛化性能。 它可以用作一個標註工具，更有效率地獲得各種對象的分割遮罩，並實現高精度的分割。

## 附錄

在3.2.2節中，我們比較了我們的無參數互動編碼方法EGD與幾種現有的編碼方法：歐幾里德距離變換、高斯熱圖和測地距離變換。 每個替代方法都依賴一個超參數。 在這裡，我們報告了找到這些方法的最佳超參數值的詳細資訊。 資料集和CNN結構與第3.2.2節所使用的相同。

8.1. 不同閾值對歐幾里德距離變換的影響

DeepIOS（Xu et al.，2016）和ISLD（Li et al.，2018）使用歐幾里德距離來編碼使用者互動。 為了有效地表示，距離變換被截斷為一個閾值。 然後，截斷的編碼圖與輸入圖像串聯在一起，引導CNN獲得分割結果。 依照這些工作的方法，我們將所有編碼圖重新縮放到[0,1]，並使用四個不同的閾值值（即0.2、0.4、0.6和0.8）進行截斷，分別使用它們的對應提示圖來引導 CNN分別進行MRI中的胎盤和CT中的脾臟分割。 表9列出了相應的定量評估結果。 我們發現胎盤的最佳閾值值為0.6，脾臟的最佳閾值值為0.4。 因此，我們在實驗中使用了這些最佳值來進行DeepIOS（Xu et al.，2016）和ISLD（Li et al.，2018）。

8.2. 不同sigma值對高斯距離變換的影響

DEXTR（Maninis et al.，2018）和DELSE（Wang et al.，2019）使用高斯距離變換來處理使用者互動。 為了產生高斯熱圖，我們需要確定每個任務的合適sigma值。 在這項工作中，我們使用了四個不同的sigma值（即3、6、9和12）來產生胎盤和脾臟分割的高斯熱圖，分別將它們的對應熱圖與輸入圖像串聯在一起 ，引導CNN獲得初始分割。 結果列在表10。 可以發現，胎盤的最佳sigma值為9，脾臟的最佳sigma值為4。 因此，我們在實驗中使用了這些最佳值來進行DEXTR（Maninis et al.，2018）和DELSE（Wang et al.，2019）。

8.3. 不同閾值對測地距離變換的影響

DeepIGeoS（Wang et al.，2018b）使用測地距離變換來編碼使用者互動。 與歐幾里德距離變換類似，我們使用不同的閾值值截斷了測地距離圖。 這些測地距離圖與輸入圖像串聯在一起，引導CNN獲得初始分割。 定量評估結果列在表11。 可以觀察到，胎盤的最佳閾值值為0.6，脾臟的最佳閾值值為0.4。 因此，我們在實驗中使用了這些值來進行DeepIGeoS（Wang et al.，2018b）的實驗。