



## پردازش هوشمند تصاویر زیست پزشکی

نیم سال اول ۰۲-۰۳

مدرس: محمدحسین رهبان

مهلت ارسال: ۱۳ دی

تمرین پنجم

- مهلت ارسال پاسخ تا ساعت ۲۳:۵۹ روز مشخص شده است.
- در طول ترم امکان ارسال با تاخیر تمرین ها بدون کسر نمره تا سقف ۱۲ روز وجود دارد. محل بارگزاری جواب تمرین ها بعد از ۴ روز بسته خواهد شد و پس از گذشت این مدت، پاسخ های ارسال شده پذیرفته نخواهند شد.
- توجه داشته باشید که نوت بوک های شما باید قابلیت باز اجرای ۱۰۰ درصد داشته باشند و در صورت نیاز به نصب یک کتابخانه یا دسترسی به یک فایل، مراحل نصب و دانلود (از یک محل عمومی) در نوت بوک وجود داشته باشد.
- هم فکری در انجام تمرین مانعی ندارد، فقط توجه داشته باشید که پاسخ تمرین حتما باید توسط خود شخص نوشته شده باشد. همچنین در صورت هم فکری در هر تمرین، در ابتدای جواب تمرین نام افرادی که با آن ها هم فکری کرده اید را حتما ذکر کنید.
- برای پاسخ به سوالات نظری در صورتی که از برگه خود عکس تهیه می کنید، حتما توجه داشته باشید که تصویر کاملا واضح و خوانا باشد. در صورتی که خوانایی کافی را نداشته باشد، تصحیح نخواهد شد.
- محل بارگذاری سوالات نظری و عملی در هر تمرین مجزا خواهد بود. به منظور بارگذاری بایستی تمارین تئوری در یک فایل زیپ با نام `SPB_Theo_hw[HW-Number]_[First-Name]_[Last-Name]_[Student-Id].zip` و تمارین عملی نیز در یک فایل مجزای زیپ با نام `SPB_Prac_hw[HW-Number]_[First-Name]_[Last-Name]_[Student-Id].zip` بارگذاری شوند.
- در صورت وجود هرگونه ابهام یا مشکل، در کوئرای درس آن مشکل را بیان کنید و از پیغام دادن مستقیم به دستیاران آموزشی خودداری کنید.

### بخش تئوری (۴۰ نمره)

#### ۱. سوال اول (۱۰ نمره)

در مورد Object Detection به سوالات زیر پاسخ دهید.

(i) نحوه ی عملکرد آشکارساز های تک مرحله ای و دو مرحله ای را بیان کنید؟ تفاوت آن ها در چیست؟ موارد استفاده ی هر کدام در حوزه ی مدیکال چه مواردی است؟

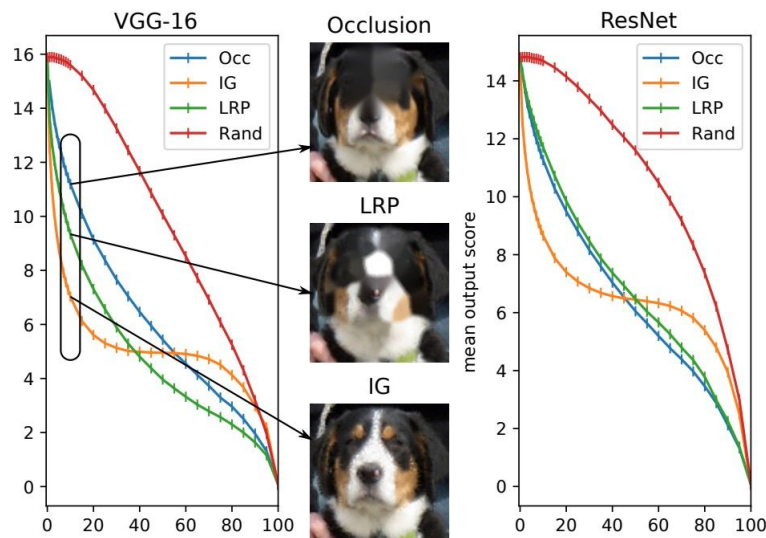
(ii) معماری R-CNN را شرح دهید. این آشکارساز چند مرحله ای عمل می کند؟ تفاوت آن با نسخه های fast R-CNN و faster R-CNN را بیان کنید؟

#### ۲. سوال دوم (۱۰ نمره)

در مورد تفسیر پذیری به سوالات زیر پاسخ دهید

(i) در attribution methods خروجی آن یعنی saliency map ها دارای نویز هستند. برای وجود این نویز سه فرضیه مطرح هست. هر کدام را توضیح دهید و اگر برای آن راه حلی وجود دارد بیان کنید.

(ii) در ارزیابی روش های attribution نحوه ی ارزیابی selectivity را توضیح دهید. برای توضیح این قسمت می توانید از الگوریتم ۳ این [survey](#) که در رابطه با pixel flipping است کمک بگیرید. همچنین به طور خاص توضیح دهید محورهای افقی و عمودی شکل زیر که در سروی آمده است نشان دهنده ی چه چیزی است؟



### ۳. سوال سوم (۲۰ نمره)

در مورد Federated Learning به سوالات زیر پاسخ دهید.

(i) تفاوت کلیدی Federated Learning با یادگیری decentralized یا distributed learning چیست؟

(ii) [مقاله ی Personalized Federated Learning with Theoretical Guarantees: A Model-Agnostic Meta-Learning Approach](#)

را مطالعه کنید. توضیح دهید در این مقاله چگونه از یک shared model به مدل های personalized برای هر کلاینت می رسد. علت ظهور مشتق مرتبه دوم (Hessian) را در الگوریتم این مقاله توضیح دهید.

---

**Algorithm 1:** The proposed Personalized FedAvg (Per-FedAvg) Algorithm

---

**Input:** Initial iterate  $w_0$ , fraction of active users  $r$ .

**for**  $k : 0$  to  $K - 1$  **do**

Server chooses a subset of users  $\mathcal{A}_k$  uniformly at random and with size  $rn$ ;

Server sends  $w_k$  to all users in  $\mathcal{A}_k$ ;

**for all**  $i \in \mathcal{A}_k$  **do**

Set  $w_{k+1,0}^i = w_k$ ;

**for**  $t : 1$  to  $\tau$  **do**

Compute the stochastic gradient  $\tilde{\nabla} f_i(w_{k+1,t-1}^i, \mathcal{D}_t^i)$  using dataset  $\mathcal{D}_t^i$ ;

Set  $\tilde{w}_{k+1,t}^i = w_{k+1,t-1}^i - \alpha \tilde{\nabla} f_i(w_{k+1,t-1}^i, \mathcal{D}_t^i)$ ;

Set  $w_{k+1,t}^i = w_{k+1,t-1}^i - \beta(I - \alpha \tilde{\nabla}^2 f_i(w_{k+1,t-1}^i, \mathcal{D}_t^i)) \tilde{\nabla} f_i(\tilde{w}_{k+1,t}^i, \mathcal{D}_t^i)$ ;

**end for**

Agent  $i$  sends  $w_{k+1,\tau}^i$  back to server;

**end for**

Server updates its model by averaging over received models:  $w_{k+1} = \frac{1}{rn} \sum_{i \in \mathcal{A}_k} w_{k+1,\tau}^i$ ;

**end for**

---

۱. سوال اول (۲۰ نمره)

در این تمرین می‌خواهیم یک مدل ناحیه‌بندی (segmentation) آموزش دهیم و به کمک آن ارگان کبد در بدن را در تصاویر CT ناحیه‌بندی کنیم. در این سوال به شما یک دیتاست داده شده است که شامل ۳۰ تصویر سه‌بعدی CT است. همانطور که می‌دانید هر تصویر سه‌بعدی از تعدادی تصویر دوبعدی تشکیل می‌شود که در محور Z روی هم قرار گرفته‌اند و تصویر حجمی را می‌سازند. به هر کدام از این تصاویر دوبعدی یک اسلایس گفته می‌شود. شکل زیر یک اسلایس از یک تصویر حجمی را نشان می‌دهد.



در این تمرین می‌خواهیم مدلی را آموزش دهیم که بتواند ارگان کبد (liver) را در یک تصویر حجمی ناحیه‌بندی کند. به طور دقیق‌تر ورودی شبکه یک تنسور با ابعاد  $num\_slices \times H \times W$  است و خروجی باید یک ماسک باینری با همین ابعاد باشد. در این تمرین دو روش برای انجام این کار پیاده می‌کنیم و نتایج دو حالت رو مقایسه می‌کنیم. برای ارزیابی از معیار دایس استفاده می‌کنیم. (در زمان ارزیابی، دایس را فقط روی اسلایس‌هایی محاسبه کنید که ارگان کبد در آنها موجود است.)

**حالت اول:** در این حالت برای ناحیه‌بندی تصاویر سه‌بعدی یک مدل دوبعدی آموزش می‌دهیم. یعنی برای آموزش این مدل یک دیتاست از اسلایس‌هایی که شامل ارگان کبد هستند می‌سازیم و با استفاده از این دیتاست یک مدل UNet دوبعدی را آموزش می‌دهیم. این مدل قادر خواهد بود که یک تصویر دوبعدی را بگیرد و در خروجی ماسک باینری ناحیه‌بندی را به ما بدهد. با استفاده از این مدل می‌توانیم تک تک اسلایس‌های یک تصویر حجمی را مستقلاً ناحیه‌بندی کنیم و به این ترتیب یک ماسک سه‌بعدی برای تصویر بدست آوریم. نوتبوک را کامل کنید و نتایج ارزیابی را گزارش دهید. (دقت کنید در پیاده‌سازی شبکه UNet از انکودر به دیکودر skip connection باید وجود داشته باشد.) شکل زیر یک نمونه معماری UNet را نشان می‌دهد.



ریاضی این معادل با گرادیان غالب‌ترین logit در آخرین activation map است. درواقع فیچرهای encode شده در آخرین لایه کانولوشنی مدل را قانع می‌کنند که آن کلاس را به عنوان کلاس نهایی انتخاب کند. برای توضیح بیشتر در مورد عملکرد این روش می‌توانید پیپر [GradCAM](#) را مطالعه کنید.

