



تحلیل هوشمند تصاویر زیست پزشکی

نیم سال اول ۰۳-۰۲

مدرس: محمدحسین رهبان

مهلت ارسال: ۲۴ آذر

پاسخ نامه

تمرین چهارم

- 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 بازاجرای 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 همفکری 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 IABI_Theo_hw4_[First-Name]_[Last-Name]_[Student-Id].zip و تمارین عملی نیز در یک فایل مجزای زیپ با نام IABI_Prac_hw4_[First-Name]_[Last-Name]_[Student-Id].zip بارگذاری شوند.
- در صورت وجود هرگونه ابهام یا مشکل، در کوثرای درس آن مشکل را بیان کنید و از پیغام دادن مستقیم به دستیاران آموزشی خودداری کنید.

بخش نظری (۳۰ نمره)

سوال اول (۱۰ نمره)

یک گراف $G = (V, E)$ را در نظر بگیرید، با ویژگی های گره $x(v)$ برای هر $v \in V$. برای هر گره $v \in V$ ، فرض کنید $h_v^{(0)} = x(v)$ نمایه اولیه گره باشد. در هر تکرار k ، نمایه ها به صورت زیر به روزرسانی می شوند:

$$h_{\mathcal{N}(v)}^{(k)} = \text{AGGREGATE}(\{h_u^{(k-1)}, \forall u \in \mathcal{N}(v)\})$$
$$h_v^{(k)} = \text{COMBINE}(h_v^{(k-1)}, h_{\mathcal{N}(v)}^{(k)}),$$

توجه کنید که آرگومان تابع $\text{AGGREGATE}(\cdot)$ ، $h_u^{(k-1)}, \forall u \in \mathcal{N}(v)$ ، یک مولتی ست است. یعنی، از آنجا که چندین گره می توانند نمایه یکسانی داشته باشند، یک عنصر می تواند چندین بار در $h_u^{(k-1)}, \forall u \in \mathcal{N}(v)$ رخ دهد. در نهایت، خود یک گراف می تواند با محاسبه برخی توابع اعمال شده بر روی مولتی ست تمام نمایه های گره در یک تکرار نهایی K ، که ما آن را به صورت زیر نمایش می دهیم، نمایه بندی شود:

$$\text{READOUT}(\{h_v^{(K)}, \forall v \in V\})$$

ما می‌خواهیم از نمایه‌های گراف فوق برای آزمایش اینکه آیا دو گراف $G_1 = (V_1, E_1)$ و $G_2 = (V_2, E_2)$ ایزومورف هستند یا خیر، استفاده کنیم. به یاد داشته باشید که این صحیح است اگر و فقط اگر یک تناظر یک به یک $\phi : V_1 \rightarrow V_2$ بین گره‌های G_1 و گره‌های G_2 وجود داشته باشد به طوری که برای هر $u, v \in V_1$

$$(u, v) \in E_1 \Leftrightarrow (\phi(u), \phi(v)) \in E_2$$

روشی که ما از مدل فوق برای آزمایش ایزومورفی استفاده می‌کنیم به شرح زیر است. برای دو گراف، اگر توابع خواندن آن‌ها متفاوت باشد، یعنی

$$\text{READOUT}(\{h_v^{(K)}, \forall v \in V_1\}) \neq \text{READOUT}(\{h_v^{(K)}, \forall v \in V_2\}),$$

ما نتیجه می‌گیریم که گراف‌ها ایزومورف نیستند. در غیر این صورت، ما نتیجه می‌گیریم که گراف‌ها ایزومورف هستند. توجه داشته باشید که این الگوریتم کامل نیست و ما به بررسی قدرت این نمایه‌های گراف خواهیم پرداخت.

۱.۰ انتخاب تابع تجمیع

انتخاب تابع $\text{AGGREGATE}(\cdot)$ برای قدرت مدل فوق مهم است. سه انتخاب رایج عبارتند از:

$$\text{AGGREGATE}_{\max}(\{h_u^{(k-1)}, \forall u \in \mathcal{N}(v)\})_i = \max_{u \in \mathcal{N}(v)} (h_u^{(k-1)})_i \text{ (حداکثر تک‌عنصری)}$$

$$\text{AGGREGATE}_{\text{mean}}(\{h_u^{(k-1)}, \forall u \in \mathcal{N}(v)\}) = \frac{1}{|\mathcal{N}(v)|} \sum_{u \in \mathcal{N}(v)} (h_u^{(k-1)})$$

$$\text{AGGREGATE}_{\text{sum}}(\{h_u^{(k-1)}, \forall u \in \mathcal{N}(v)\}) = \sum_{u \in \mathcal{N}(v)} (h_u^{(k-1)})$$

مثالی از دو گراف $G_1 = (V_1, E_1)$ و $G_2 = (V_2, E_2)$ و ویژگی‌های اولیه گره‌های آن‌ها ارائه دهید، به گونه‌ای که برای برخی گره $v_1 \in V_1$ و برخی گره $v_2 \in V_2$ با ویژگی‌های اولیه یکسان $h_{v_1}^{(0)} = h_{v_2}^{(0)}$ ، ویژگی‌های به‌روزرسانی شده $h_{v_1}^{(1)}$ و $h_{v_2}^{(1)}$ در صورت استفاده از تجمیع میانگین و حداکثر یکسان باشند، اما در صورت استفاده از تجمیع جمع متفاوت باشند.

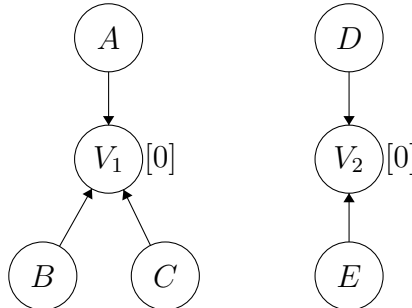
راهنمایی: ویژگی‌های گره‌های شما می‌توانند اسکالر به جای بردار باشند، یعنی ویژگی‌های گره تک‌بعدی به جای چندبعدی. همچنین، شما آزاد هستید تا تعداد گره‌ها (مثلاً ۳ گره)، اتصالات آن‌ها (یعنی یال‌های بین گره‌ها) را در مثال خود به دلخواه انتخاب کنید.

پاسخ:

فرض کنید دو گراف G_1 و G_2 دارای دو گره با ویژگی‌های اولیه برابر هستند، به طوری که

$$h_{v_1}^{(0)} = h_{v_2}^{(0)}.$$

برای این دو گراف، تابع تجمعی mean و max ویژگی‌های به‌روزرسانی شده یکسان تولید می‌کند، اما تابع تجمعی sum ویژگی‌های متفاوتی تولید می‌کند. گراف‌ها و ویژگی‌های آن‌ها به شکل زیر هستند:



ویژگی‌های اولیه گره‌ها هر دو گراف به صورت زیر هستند:

$$h_A^{(0)} = h_D^{(0)} = 20, h_B^{(0)} = h_C^{(0)} = 5, h_E^{(0)} = 0$$

با استفاده از توابع تجمیعی مختلف، ویژگی‌های به‌روزرسانی شده برای گره‌های v_1 و v_2 به صورت زیر خواهند بود.

• تجمیع میانگین

$$h_{v_1}^{(1)} = \frac{1}{3}(20 + 5 + 5) = 10 = h_{v_2}^{(1)} = \frac{1}{2}(20 + 0) = 10$$

• تجمیع حداکثر

$$h_{v_1}^{(1)} = \max(20, 5, 5) = 20 = h_{v_2}^{(1)} = \max(20, 0) = 20$$

• تجمیع جمع

$$h_{v_1}^{(1)} = 20 + 5 + 5 = 30 \neq h_{v_2}^{(1)} = 20 + 0 = 20$$

در نتیجه، برای گره v_1 در گراف G_1 و گره v_2 در گراف G_2 ، توابع تجمیعی mean و max ویژگی‌های یکسانی تولید می‌کنند، اما تابع تجمیعی sum ویژگی‌های متفاوتی تولید می‌کند.

۲۰۰ آزمون وایسفایلر-لهمن

الگوریتم آزمایش ایزومورفی ما حداکثر به اندازه آزمون معروف وایسفایلر-لهمن (WL-Test) قدرتمند است. در هر تکرار، این الگوریتم نمایه هر گره را به مجموعه‌ای که شامل نمایه قبلی خود گره و نمایه‌های قبلی تمام همسایه‌هایش است، به‌روزرسانی می‌کند. الگوریتم کامل در شکل زیر آمده است.

Algorithm 3: Weisfeiler-Lehman Test

Data: $G_1 = (V_1, E_1)$, $G_2 = (V_2, E_2)$, initial features $x(\cdot)$, number of iterations K

Result: Prediction of whether G_1 and G_2 are isomorphic

for $v \in V_1 \cup V_2$ **do**

$l_v^{(0)} \leftarrow x(v)$

end

for $k = 1, \dots, K$ **do**

for $v \in V_1 \cup V_2$ **do**

$l_v^{(k)} \leftarrow \text{HASH}(l_v^{(k-1)}, \{l_u^{(k-1)}, \forall u \in \mathcal{N}(v)\})$

end

end

return $\{l_v^{(K)}, \forall v \in V_1\} = \{l_v^{(K)}, \forall v \in V_2\}$

ثابت کنید که مدل عصبی ما حداکثر به اندازه آزمون WL قدرتمند است. به طور دقیق‌تر، فرض کنید G_1 و G_2 غیر ایزومورف هستند و فرض کنید که نمایه‌های گره‌های آن‌ها در طی K تکرار با همان توابع $\text{AGGREGATE}(\cdot)$ و $\text{COMBINE}(\cdot)$ به‌روزرسانی می‌شوند. نشان دهید که اگر رابطه زیر برقرار باشد آنگاه آزمون WL نیز تصمیم می‌گیرد که گراف‌ها ایزومورف نیستند.

$$\text{READOUT}(\{h_v^{(K)}, \forall v \in V_1\}) \neq \text{READOUT}(\{h_v^{(K)}, \forall v \in V_2\})$$

توجه: اثبات باید برای هر تابع READOUT، COMBINE، AGGREGATE عمومی باشد. به عبارت دیگر، کافی نیست که این موضوع را فقط برای یک نمونه خاص از مدل GNN نشان دهید. **راهنمایی:** می‌توانید از اثبات توسط تناقض استفاده کنید با این فرض که آزمون وایسفالر-لهمن نمی‌تواند در پایان تکرار K تصمیم بگیرد که آیا G_1 و G_2 ایزومورف هستند یا خیر.

پاسخ:

فرض کنید دو گراف $G_1 = (V_1, E_1)$ و $G_2 = (V_2, E_2)$ داریم که ایزومورف نیستند. ویژگی‌های گره‌های آنها را در نظر بگیرید که بیش از K تکرار با همان توابع $AGGREGATE(\cdot)$ و $COMBINE(\cdot)$ به‌روزرسانی شده‌اند.

برای رسیدن به تناقض، فرض کنید که آزمون وایسفالر-لهمن (WL) می‌گوید که G_1 و G_2 پس از K تکرار ایزومورف هستند. این بدان معناست که برای هر گره $v_1 \in V_1$ ، گره متناظری $v_2 \in V_2$ وجود دارد به طوری که برچسب‌های نهایی (ویژگی‌های نهایی) اختصاص داده شده توسط آزمون WL یکسان هستند، یعنی

$$l_{v_1}^{(K)} = l_{v_2}^{(K)}.$$

با این حال، تابع READOUT مدل عصبی برای دو گراف نتایج متفاوتی به دست می‌دهد، یعنی:

$$READOUT(\{h_v^{(K)}, \forall v \in V_1\}) \neq READOUT(\{h_v^{(K)}, \forall v \in V_2\}).$$

این بدان معنی است که حداقل برای یک جفت گره، مثلاً $v_1 \in V_1$ و $v_2 \in V_2$ ، ویژگی‌های متفاوتی وجود دارد:

$$h_{v_1}^{(K)} \neq h_{v_2}^{(K)}.$$

با توجه به اینکه ویژگی‌ها با استفاده از همان توابع AGGREGATE و COMBINE محاسبه شده‌اند، این مستقیماً با نتیجه‌گیری آزمون WL که G_1 و G_2 ایزومورف هستند، در تناقض است. بنابراین، اگر تابع READOUT مدل عصبی تفاوتی را بین دو گراف تشخیص دهد، آزمون WL نیز باید تفاوتی را تشخیص دهد، که نشان می‌دهد مدل عصبی ما حداکثر به اندازه آزمون WL قدرتمند است.

سوال دوم (۱۰ نمره)

مقاله [Hierarchical Graph Representations in Digital Pathology](#) را بخوانید و به سوالات زیر پاسخ دهید.

۱. رویکرد نوآورانه‌ای که در مقاله در مورد استفاده از نمایش‌های گرافی سلسله مراتبی در پاتولوژی دیجیتال معرفی شده چیست و چگونه از روش‌های پیشین در تحلیل تصاویر پاتولوژیک متفاوت است؟

پاسخ:

رویکرد جدید معرفی شده در مقاله، استفاده از نمایش‌های گرافی سلسله مراتبی چند سطحی برای نمونه‌های بافتی است. این روش با در نظر گرفتن بافت به عنوان ترکیبی سلسله مراتبی از انواع مختلف موجودیت‌های هیستولوژیکی، از روش‌های سنتی متفاوت است. در حالی که روش‌های سنتی بر تحلیل مستقیم تصاویر از طریق تکنیک‌های پردازش تصویر و یادگیری ماشین تمرکز دارند، این رویکرد جدید اطلاعات بافتی را در سطوح مختلف و با توجه به تعاملات و روابط بین انواع مختلف موجودیت‌های هیستولوژیکی encode می‌کند.

البته این صرفاً برداشت من از مقاله است و جواب‌های شما با توجه به کلیت خود مقاله بررسی می‌شوند.

۲. در زمینه نمایش‌های گرافی سلول-به-بافت سلسله مراتبی (HACT) که در مقاله استفاده شده است، شبکه عصبی گرافی سلسله مراتبی (HACT-Net) چگونه تصاویر هیستولوژی ورودی را برای طبقه‌بندی ساختارهای بافتی را پردازش می‌کند؟

پاسخ:

شبکه معرفی شده HACT-NET از دو زیرشبکه CG-GNN و TC-GNN ساخته شده است. این ۲ زیر شبکه بازنمایی‌هایی در لایه‌های سلولی و بافتی ایجاد میکنند که با استفاده از آنها میتوانیم مسائل طبقه‌بندی را با دقت بیشتری به جزئیات تصویر حل کنیم. در واقع این مقاله برای بررسی بهتر ویژگی‌های local و از همچنین دسترسی به الگوهای global دست به استفاده از همزمان از مدل‌های GNN در دو سطح متفاوت و سپس ترکیب آنها برحسب همسایگی‌شان می‌زنند. البته این صرفاً برداشت من از مقاله است و جواب‌های شما با توجه به کلیت خود مقاله بررسی می‌شوند.

مقاله *SlideGraph⁺* در کلاس معرفی شده است. با توجه به آن به سوالات زیر پاسخ دهید.

۱. چگونه *SlideGraph⁺* از مفهوم پیش‌بینی‌های سطح گره در معماری شبکه خود استفاده می‌کند و این رویکرد چه مزایایی نسبت به روش‌های سنتی تحلیل تصویر کلی اسلاید (WSI) دارد؟

پاسخ:

در *SlideGraph⁺*، هر گره نمایانگر ناحیه یا جنبه خاصی از ساختار سلولی در WSI است. تمرکز روی گره‌ها بجای کل تصویر به *SlideGraph⁺* این امکان را می‌دهد تا پیش‌بینی‌های دقیق‌تر و محلی‌تری در مورد ویژگی‌های بافت ارائه دهد. این سطح از جزئیات به درک دقیق‌تر ساختار میکروسکوپی بافت و ارتباط آن با وضعیت HER2 کمک می‌کند. البته این صرفاً برداشت من از مقاله است و جواب‌های شما با توجه به کلیت خود مقاله بررسی می‌شوند.

۲. نقش مدل رگرسیون چگالی DAB در *SlideGraph⁺* چیست و چگونه این مدل پیش‌بینی وضعیت HER2 از تصاویر رنگ‌آمیزی شده H&E را بهبود می‌بخشد؟

پاسخ:

این مدل، که شدت DAB را مستقیماً از تصاویر H&E پیش‌بینی می‌کند، نیاز به رنگ‌آمیزی IHC اضافی را دور می‌زند و فرآیند تشخیص را ساده‌تر می‌کند. همچنین استفاده از آن نه تنها باعث صرفه‌جویی در زمان و منابع می‌شود، بلکه باعث می‌شود دقت بالایی در پیش‌بینی وضعیت HER2 داشته باشیم. البته این صرفاً برداشت من از مقاله است و جواب‌های شما با توجه به کلیت خود مقاله بررسی می‌شوند.

منابع پیشنهادی برای مطالعه بیشتر

- [Graph Neural Networks and Their Current Applications in Bioinformatics](#)
- [Representation Learning of Histopathology Images Using Graph Neural Networks](#)
- [Visualization for Histopathology Images using Graph Convolutional Neural Networks](#)

سوال سوم (۱۰ نمره)

مدل VoxMorph علی‌رغم پیشرفت‌هایش، در registration تصاویر پزشکی محدودیت‌هایی دارد. از جمله آن‌ها می‌توان به محدودیت آن در یادگیری ویژگی‌های سطح بالا و ارائه‌هایی deformation با حفظ توپولوژی اشاره کرد. در سال‌های اخیر مدل‌های مبتنی بر ترنسفورمر برای وظیفه رجیستریشن تصاویر پزشکی ارائه شده‌اند، از جمله:

- **ViT-V-Net: Vision Transformer for Unsupervised Volumetric Medical Image Registration**
- **TransMorph: Transformer for Unsupervised Medical Image Registration**

با بررسی این دو مقاله، به سوالات زیر پاسخ دهید:

- هر کدام چه محدودیتی از VoxMorph را حل کرده‌اند؟ راه حل ارائه شده آن‌ها را مختصراً توضیح دهید.
- معماری این دو مدل را بررسی کرده و شباهت‌ها و تفاوت‌های آن‌ها با مدل VoxMorph را بیان کنید.

پاسخ:

TransMorph محدودیت ConvNets در VoxMorph را با ترکیب معماری‌های ترانسفورماتور، که روابط فضایی دوربرد را در تصاویر بهتر مدیریت می‌کند، برطرف می‌کند. ReceptiveField بزرگتر آن امکان درک دقیق‌تری از مکاتبات فضایی را فراهم می‌کند. TransMorph یک مدل ترکیبی است که ترانسفورماتورها و ConvNet را ترکیب می‌کند. این شامل انواع دیفیومورفیک و بیزی است که از تغییر شکل‌های صحیح توپولوژیکی اطمینان می‌دهد و تخمین‌های عدم قطعیت کالبره‌شده را در ثبت ارائه می‌دهد.

ViT-V-Net با ادغام Vision Transformer (ViT) با ConvNets، به محدودیت ConvNet‌ها در ثبت روابط فضایی دوربرد می‌پردازد، که به تمرکز بر روی بخش‌های مرتبط تصویر برای ثبت بهتر کمک می‌کند. ViT-V-Net و ConvNet را برای بهبود محلی‌سازی دقیق، که برای ثبت تصویر بسیار مهم است، پل می‌کند. این ابزار از مکانیسم خودتوجهی ViT برای یادگیری روابط فضایی دوربرد استفاده می‌کند و قابلیت‌های ConvNet را تکمیل می‌کند.