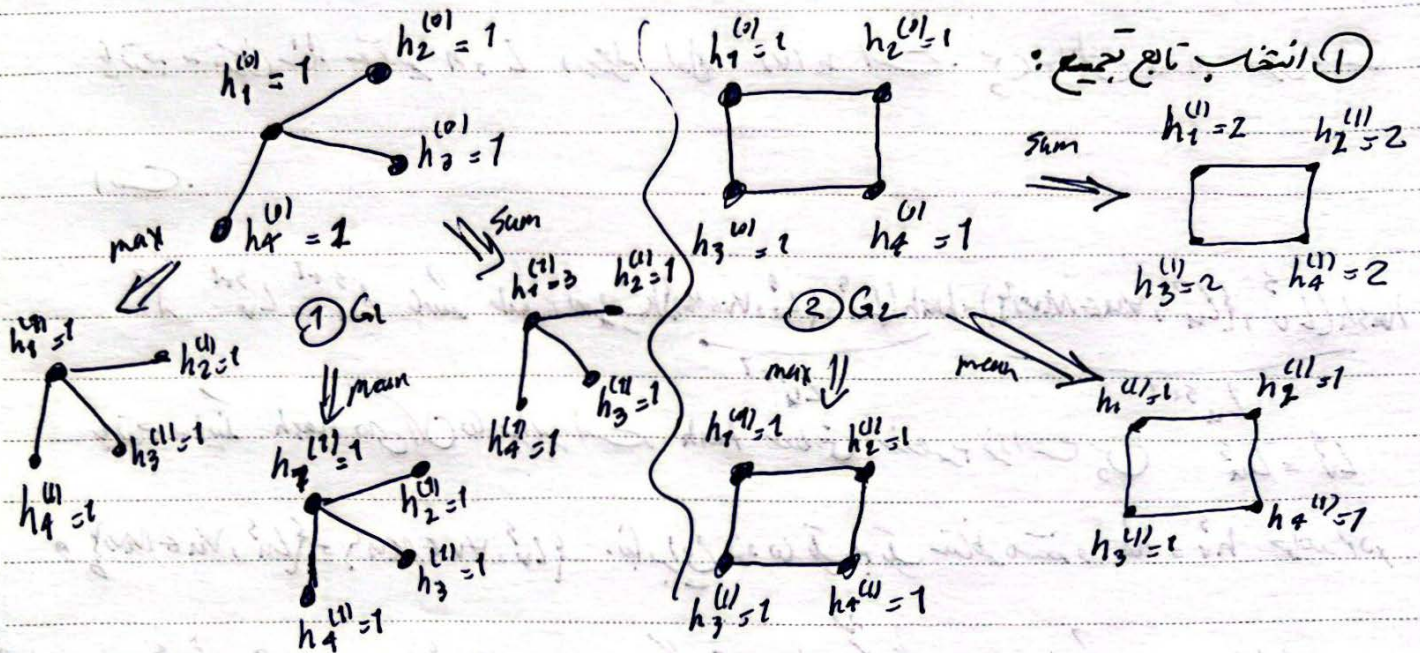


سوال اول: همگرایی با اینر محدوده، روش زمانی تبار



تابع‌های \max, mean برخلاف تابع sum به مقدار ورودی‌ها وابسته نیستند. بنابراین در حالت \max یا mean برای گره‌های دوگراف مانند گراف‌های بالا (با تعداد گره مساوی) مقادیر \max, mean و در نتیجه مقدار $h_i^{(t)}$ همانند حالت قبل خواهد بود. اما برای تابع sum اینگونه نیست.

② لایه‌های داینامیک-لایه‌های ایستاد با استفاده از بهای تلف می‌فرستند
 به W_L مصحح بکریه گراف‌ها
 این روش استفاده از
 فرم تلف زمانی به قرار است که دو مجموعه $\{v_1, v_2, \dots, v_n\}$ و $\{v_1, v_2, \dots, v_n\}$ برابر باشند.
 اگر این فرم به برابر بود ReLU می‌گیریم، به تناقض خواهیم رسید.

بارتجه به اینکه L ها نتیجه Mush هستند، information loss کمتری دارند. بنابراین ثابت کنیم برابر لایه‌های دوگراف به برابر h های لایه‌های نتیجه می‌رسد.

Subject :

Year .

Month .

Date .

()

اثبات با استقرا : به ازای هر $K > m$ داریم اگر $L_u^m = L_v^m$ باشد، آنگاه $h_u^m = h_v^m \Rightarrow$ فرضی

پایه به فرضی شده معکوس h و L در مرحله اول $n(u)$ است. پس به ازای $m=0$ فرضی درست

است.

اگر $L_u^{s+1} = L_v^{s+1}$ باشد طبق تعریف $hash(L_u^s, \{L_u^s, v_u \in M(u)\}) = hash(L_v^s, \{L_v^s, v_v \in M(v)\})$

پایه به ایند $hash$ می‌کند و برابر است $hash$ مقادیر مشخصه فرد است پس $L_v^s = L_u^s$

و $\{L_u^s, v_u \in M(u)\} = \{L_v^s, v_v \in M(v)\}$. بنابراین نزد و یک به یک متناظر هستند و $h_v^s = h_u^s$ خواهد بود.

$\{h_v^s, v_v \in M(v)\} = \{h_u^s, v_u \in M(u)\}$. بنابراین اگر معکوس h های دیگران برابر باشد، حاصل مقدار تابع

Aggregation کل ها نیز برابر و در نتیجه حاصل تابع Combine کل ها نیز برابر خواهد بود.

در نتیجه $random$ کل ها نیز یکسان خواهد بود که متناقض با فرضی مسئله است.

بنابراین حکم اثبات می شود.

Hierarchical Graph Representation in Digital Pathology :

سؤال دوم :

① در یکدانه‌دانه ارائه شده ، توسعه یک *entity-graph* سلسله مراتبی چندسطحی (Hac) ، برای یک شبکه عصبی گرافی سلسله مراتبی (Hac-Net) برای تفسیر و تحلیل بافت‌شناسی است . این رویکرد امکان encode کردن جامع *microenvironment* های سلول و بافت را فراهم می‌کند و غنای دقیق و جامع تری از ترکیب بافت ارائه می‌دهد . این روش با گرفتن ترکیبات سلسله مراتبی موجود در بافت‌شناسی و تعاملات داخلی و خارجی کان ها ، درک جامع تری از ساختار بافت ارائه می‌دهد که برای تشخیص دقیق سرطان بسیار مهم است .

اکثر روش‌های موجود ، پس از شبکه‌های عصبی مخزن‌گرفته به شکل *patch-wise* استفاده می‌کنند که اطلاعات مربوط به محل بافت را در نظر نمی‌گیرد ، فقط اطلاعات *patch-level* محلی را مورد استفاده قرار می‌دهد . همچنین روش‌های قبلی معمولاً از یک *single-stream* در طبقه‌بندی تصاویر بافت‌شناسی استفاده می‌کنند . در مقابل ، در روش ارائه شده از یک غنای گراف موجود در سلسله مراتبی (Hac) استفاده می‌کنند که اطلاعات مربوط به سلول و بافت را encode می‌کند و همچنین از شبکه عصبی سلسله مراتبی گرافی استفاده می‌کند . در نتیجه این کار ورودی غنای گرافی را به صورت سلسله مراتبی پردازش می‌کنند . بنابراین رویکرد جامع تری برای تفسیر و تحلیل ترکیب بافت ارائه می‌دهد .

② Hac-Net تصاویر بافت‌شناسی را با استفاده از یک شبکه عصبی گرافی سلسله مراتبی برای طبقه‌بندی پردازش می‌کند . این مدل به شیوه سلسله مراتبی به روش غنای گراف سلسله مراتبی سلول و بافت (Cell to tissue Hac) از سطح سلول تا سطح ناحیه بافت در دست عمل می‌کند . هدف این شبکه بدست آوردن یک *contextualized graph embedding* است که توزیع مورفولوژی و تولوژی موجود در بافت را encode می‌کند . پس این *embedding* برای طبقه‌بندی گراف ورودی

استفاده می شود به شکلی که Node - level اجازه می دهد تا به طور مؤثر تصاویر بافت شامی را بر اساس نمای های

گراف سلسله مراتبی تجزیه و تحلیل و طبقه بندی کند.

Slide Graph⁺ :

① در مدل Slide Graph⁺ ، embedding ویژگی های نودهای موجود در شبکه عصبی گرافی به یک لایه خطی متناظر با پیش بینی سطح node را ایجاد کند. این prediction score های node-level برای ایجاد پیش بینی wsol-level با هم ترکیب می شوند.

روشی های منطقی مبتنی بر Patch ، ندهای محلی که یک یک تصویر را تحلیل می کنند و ممکن است ساختار کلی بافت را در نظر بگیرند. در بسیاری از موارد فقط برچسب مربوط به wsol-level در دسترس است. Slide Graph⁺ این چالش را با مدل سازی

ارتباط در Patch با یک label خاص برطرف می کند. در این روش با ایجاد پیش بینی node-level ، نه تنها score مربوط به

prediction کلی را مشخص می کند، بلکه گروه های فعلی گراف را نیز شناسایی می کند که منجر به score کلی تصویر شده اند. این

موضوع مدل را تقسیم نموده و بررسی می کند و پیش بینی هایی در مورد مناطق خاص که روی پیش بینی تأثیر می گذارند ارائه می دهد.

② مدل را بر اساس چگالی DAB در Slide Graph⁺ و وضعیت HER2 را با تعیین چگالی رنگ لکسیری در تصاویر رنگ لکسیری

شده با H&E پیش بینی می کند (بررسی شدت DAB staining) این موضوع بسیار مهم است زیرا مدل می تواند با

بررسی شدت DAB intensity ، سطح HER2 را تعیین کرده و وضعیت کان را در نمونه بافت مشخص کند و این

کار را به صورت مستقیم و بدون نیاز به رنگ لکسیری ایمنو هستو شیمی اضافی (IHC) انجام می دهد و نیاز به بافتور به کان را از بین

می برد. به طور معمول IHC برای ارزیابی وجود گیرنده های پرده ای خاص مانند HER2 انجام می شود. این مدل کارایی و

دقت پیش بینی در وضعیت HER2 را بهبود می بخشد.

سوال سوم:

ViT-V-Net:

یکی از محدودیت‌های Voxel Morph این است که به شبکه‌های کانولوشنی محلی متکی است و ممکن است روابط long-range spatial در تصاویر را به طور کامل ثبت نکند. (برای مثال در Voxel که لازم باشد شبکه‌ی دارند) و این منجر به دلیل local بودن عملیات مربوط به کانولوشن است. مدل ارائه شده در این مقاله با استفاده از یک ساختار ترکیبی از ConvNet و Transformer محدودیت مربوط به Voxel Morph حل می‌کند. در این روش ViT به بررسی ویژگی‌های high level اعمال شده تا مدل را ملزم به یادگیری روابط long-range در تصویر کند.

به جای استفاده از U-Net ساده،

encoder, decoder

هر دو مدل از ساختار کانولوشنی () استفاده می‌کنند با این تفاوت که در مدل ViT-V-Net، (encoder) نیاز تبدیل داده ورودی به ویژگی‌های high-level توسط لایه‌های کانولوشنی، از مدل ViT استفاده شده که در آن

این ویژگی‌ها را به vector های تبدیل و سپس به لایه latent تحولی می‌دهد. به هر کدام از embed های patch ها

positional embeds ارائه شده و سپس با استفاده از یک Transformer encoder لایه‌های Multi-head self attention

Multi-layer preception، encode شده، به از reshape کردن خروجی ViT توسط یک V-Net decoder

decode می‌شود. همچنین در مدل از spatial transformation استفاده شده است تا تصویر warp

تکثیر و برای اسکن سه بعدی بکشد. در محاسبات مدل نیز از skip connection بهره گرفته شده است.

در این مدل نیز stack شده تصویر moving و fixed است.

TransMorph: مدل‌های کانولوشن در مدل سازی روابط long range spatial (رابطه میان voxels که

ازهم دور هستند) در یک تصویر محدودیت‌هایی به دلیل locality عملیات کانولوشن دارند. معماری‌های U-Net

و V-Net برای حل این مشکل ارائه شدند که با down sampling و up sampling که از نظر تئوری باعث بروز

مشکل receptive field در ConvNet‌های بزرگ، سعی در رفع این محدودیت داشته. بنابراین حل چند مشکل باقی می‌ماند.

اول اینکه، receptive field لایه‌های اول همچون درست اندازه kernel محدود شده اند و اطلاعات global

تصویر فقط در لایه‌های محلی ترمقابل شده است. دوم اینکه با عمیق شدن لایه‌های ConvNet، تأثیر voxel

هایی که دراز یکدیگر هستند کاهش می‌یابد. بنابراین لایه‌های receptive field U-Net در عمل یار کمتری

حالت نهایی لکن است. همین موضوع ب محدودیت U-Net در نتایج VoxMorph در درک اطلاعات semantic

در مدل سازی روابط long-range بین نقاط می‌گردد. در مقابل transformer ها به دلیل داشتن receptive field

بزرگ‌تر و مکانیزم self-attention، توانایی حل این مشکل را دارند. مدل ارائه شده در این مقاله نیز از یک

ترکیب Transformer-ConvNet، با نام TransMorph استفاده می‌کند که از یک encoder برپای

swin Transformer برای ثبت spatial correspondence میان تصویر moving، fixed استفاده می‌کند.

سوی یک decoder برپای ConvNet اطلاعات خروجی transformer را decode، تبدیل لای

نبشی دهد. همچنین لازم skip connection نیز برای حفظ جریان اطلاعاتی localization میان encoder، decoder

استفاده شده است. در این مقاله ارائه شده که مدل ارائه شده به positional encoding نیاز ندارد و مدل ارائه شده

از receptive field بزرگ‌تری نسبت به ConvNet بهره استفاده در VoxMorph و U-Net برخوردار است.

Subject :

Year . Month . Date . ()

هر دو مدل از ساختار کنونی مشکل از encoder, decoder استفاده می کنند، ورودی کل ما زوج یکس

moving, fixed است با این تفاوت که در مدل TransMorph، از یک مدل ترکیبی از ساختار

Transformer, ConvNet استفاده می کنند، در نتیجه استفاده از مدل Swin Transformer محال در

یافتن روابط long range بهبود میابد.