هفته ملی پژوهش و فناوری

نمایشگاه هفته پژوهش و فناوری
دانشگاه صنعتی همدان
آذر ۱۴۰۲

# مروری بر روش‌های دسته‌بندی تصاویر هیستوپاتولوژی سینه

مرجان نمکی، مجید اسدی

گروه کامپیوتر، دانشگاه صنعتی همدان
asadi@hut.ac.ir

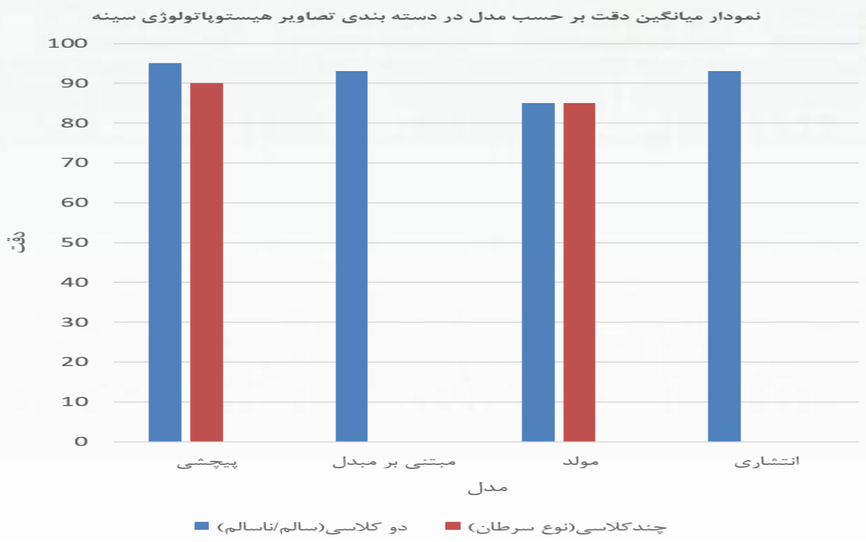
## چکیده

در این مطالعه، برخی از کارهای مهم اخیر در زمینه دسته‌بندی تصاویر بافت‌آسیب‌شناسی سرطان سینه با بهره‌گیری از روش‌های یادگیری عمیق مورد بررسی قرار گرفته است. هدف از این کار ارائه یک جمع‌بندی از روش‌ها و فناوری‌های استفاده شده در پیش‌بینی سرطان سینه و ارائه راهبردهای پیشنهادی به منظور رفع نقایص و چالش‌های موجود در این حوزه می‌باشد. روش‌های مورد مطالعه به چهار دسته مدل‌های مبتنی بر شبکه‌های عصبی پیچشی، مدل‌های مبتنی بر مبدل‌ها، مدل‌های مولد و مدل‌های انتشاری تقسیم می‌شود. هر کدام از این دسته‌ها سعی بر ارائه راه‌حل برای مسائل دوکلاسی (سالم در مقابل بیمار سرطانی) یا چندکلاسی (تشخیص درجه سرطان) دارند. در هر دسته، میانگین دقت برای مسایل دوکلاسی و چندکلاسی در یک چارت به طور خلاصه جمع‌آوری شده است. به علاوه، چند تا از مهم‌ترین چالش‌ها و نیازهایی که برای ادامه پژوهش در این زمینه وجود دارد معرفی شده است که از جمله مهم‌ترین آن‌ها می‌توان به توسعه الگوریتم‌هایی اشاره کرد که نتایج ارزیابی قابل قبول مستقل از مجموعه دادگان داشته باشند. همچنین کاهش قابل توجه پارامترها در شبکه‌های عصبی عمیق و طراحی شبکه‌های با محاسبات کم‌تر از دیگر نیازهای این زمینه است.

**واژه‌های کلیدی:** تصاویر هیستوپاتولوژی، سرطان سینه، شبکه عصبی عمیق.

## مقدمه

سرطان سینه کشنده‌ترین نوع سرطان در خانم‌هاست که تشخیص زودهنگام آن از بسیاری از صدمات ناشی از آن جلوگیری کرده و روند درمان را تسریع می‌سازد. یکی از روش‌های تشخیصی، بررسی تصاویر هیستوپاتولوژی (بافت‌آسیب‌شناسی) است که پس از تصاویر سونوگرافی و ماموگرافی به منظور تمایز نوع سرطان صورت می‌گیرد [۱]. با بهره‌گیری از روش‌های به‌روز در حوزه پردازش تصویر و کشف الگوهای موجود در آن‌ها، می‌توان تصاویر هیستوپاتولوژی سینه را در کمترین زمان و بالاترین دقت ممکن دسته‌بندی کرد و در اختیار هیستوپاتولوژیست به منظور کمک در فرایند تصمیم‌گیری، قرار داد [۲-۳]. با وجود نقش مهم روش‌های یادگیری عمیق همچنان چالش‌های زیادی در این زمینه وجود دارد.




**شکل ۱. نمودار میانگین دقت بر حسب مدل. در مطالعات صورت گرفته برای دو مدل انتشاری و مبتنی بر مبدل اطلاعات مبنی بر دسته‌بندی چندکلاسی یافت نگردید، لذا برای این دو مدل نمودار دقت چندکلاسی (دسته‌بندی براساس نوع سرطان) رسم نشده است.**

## اهداف و روش پژوهش

مقالات ارزیابی شده در این پژوهش، در سال‌های ۲۰۱۵ تا ۲۰۲۳ مورد بررسی واقع شده‌اند. اهداف این پژوهش عبارتند از: ۱) تعیین جایگاه روش‌های یادگیری عمیق در تجزیه و تحلیل تصاویر پزشکی، ۲) کاربرد یادگیری عمیق در تجزیه و تحلیل تصاویر هیستوپاتولوژی سرطان سینه، ۳) اطلاع از روش‌ها و تکنولوژی‌هایی که اخیرا در این حوزه مورد استفاده واقع شده‌اند، ۴) شناخت چالش‌ها و نقاط ضعفی که می‌توان با کمک آن‌ها مسیر تحقیقات آتی را معین نمود.

## یافته‌های پژوهش

آنچه از مطالعه مقالات اخیر به دست می‌آید این است که محققان در ابتدا به دنبال افزایش دقت در دسته‌بندی تصاویر به گروه سالم و سرطانی و در صورت سرطانی بودن، تعیین نوع آن بوده‌اند و در ادامه با دستیابی به دقت‌های بالا به دنبال بهبود روش‌ها، رفع نقایص موجود و بهینه‌سازی مدل‌ها از نظر کاهش تعداد پارامترها، افزایش سرعت محاسبات و مواردی از این دست بودند. از نظر اجرایی، تحقیقات صورت گرفته‌ی مبتنی بر شبکه‌های عصبی عمیق به چهار دسته اصلی: روش‌های مبتنی بر شبکه‌های عصبی پیچشی (CNN)، روش‌های مبتنی بر مبدل (Transformer)، مدل‌های مولد (Generative) و مدل‌های انتشاری (Diffusion) تقسیم می‌شوند. یافته‌های پژوهش حاکی از آن است که استفاده از مدل‌های پیچشی بیشترین سهم را در این زمینه تحقیقاتی به خود اختصاص داده و در رتبه‌های بعد ترنسفورمرها، مدل‌های مولد و نهایتا مدل‌های انتشاری قرار می‌گیرند. معماری‌های متعددی در روش‌های مبتنی بر CNN استفاده شده است که از جمله آنها می‌توان به LeNet5، AlexNet، VGGNet، ResNet، Inception، MobileNet، DenseNet، EfficientNet و YOLO اشاره کرد. بهره‌گیری از مدل‌های مولد با وجود اینکه توانست دقت بالایی را در مدل‌ها ایجاد کند تفاوت واضح تصاویر تولیدشده با تصاویر واقعی استفاده از این مدل‌ها را در کاربردهای پزشکی با شکست مواجه کرد. مدل‌های انتشاری جدیدترین متد مورد استفاده محققان است که تنها در چند پژوهش پزشکی و به طور خاص در یک پژوهش حوزه تشخیص سرطان سینه به کار گرفته شده است. گرچه اخیرا مدل‌های یادگیری عمیق متعددی، با بهره‌گیری از روش‌های بینایی ماشین و شناسایی الگوهای موجود در تصاویر، عملیات دسته‌بندی تصاویر هیستوپاتولوژی سینه را با دقت‌های بالا، به خوبی انجام داده‌اند، همچنان چالش‌هایی در این حوزه وجود دارد. مواردی چون عدم بهره‌گیری کامل مدل‌ها از ویژگی‌های موجود در تصاویر بافت‌آسیب‌شناسی، وجود مجموعه دادگان متعدد و عدم بررسی‌های جامع این مجموعه داده‌ها و مواردی از این دست، از جمله چالش‌های موجود در این زمینه هستند. همچنین علیرغم این که مجموعه دادگان پزشکی نسبت به دادگان عمومی برای بررسی و تست مدل‌های جامع‌تر، کوچک هستند، هنوز از مدل‌های انتشاری به صورت گسترده در این حوزه بهره گرفته نشده است.

## راهبردهای پیشنهادی

در این مطالعه سعی بر آن شد تا با گردآوری و بررسی تحقیقات صورت گرفته، تحلیلی بر روش‌های ارائه شده تا کنون صورت پذیرد. ماحصل تحقیقات شناخت چالش‌هایی بود که امکان کار بیشتر را دارد. از راهبردهای پیشنهادی می‌توان به موارد زیر اشاره کرد: ترکیب چند مجموعه داده نسبتا کوچک، اعمال مدل‌های انتشاری بر روی مجموعه دادگان کوچک و سپس آموزش یک شبکه عصبی عمیق برای دسته‌بندی تصاویر موجود در مجموعه دادگان جدید، طراحی شبکه‌های عصبی عمیق با پارامترهای کمتر، ارزیابی مستقل از مجموعه دادگان الگوریتم‌های موجود.

## منابع

[1] Ahmad, H. M., et al. (2019). Classification of breast cancer histology images using transfer learning. 2019 16th International Bhurban conference on applied sciences and technology (IBCAST), IEEE.
[2] Chhikara, B. S. and K. Parang (2023). "Global Cancer Statistics 2022: the trends projection analysis." Chemical Biology Letters 10(1): 451-451.
[3] Elston CW, Ellis IO. (1991). "The value of histological grade in breast cancer: experience from a large study with long-term follow-up." Histopathology 19: 403-410.

## تایید استاد راهنما

نام و امضا استاد راهنما: مجید اسدی

تایید تحصیلات تکمیلی:

تایید امور پژوهشی: