رده‌بندی تصاویر سرطان پوست با استفاده از یادگیری عمیق جمعی دو سطحی

پویان جابری[۱]، شهلا نعمتی*[۲]، محمد احسان بصیری[۳]

[۱] کارشناسی ارشد مهندسی کامپیوتر، دانشکدهٔ فنی و مهندسی، دانشگاه شهرکرد، شهرکرد، ایران

[۲]* استادیار گروه مهندسی کامپیوتر، دانشکدهٔ فنی و مهندسی، دانشگاه شهرکرد، شهرکرد، ایران

[۳] دانشیار گروه مهندسی کامپیوتر، دانشکدهٔ فنی و مهندسی، دانشگاه شهرکرد، شهرکرد، ایران

چکیده

برخی از مشکلات پوستی خوش‌خیم و بی‌ضرر و برخی دیگر توده‌های بدخیمی هستند که با تشخیص به‌موقع می‌توانند همچنان بی‌ضرر بمانند. در این پژوهش، یک روش یادگیری عمیق جمعی دوسطحی برای رده‌بندی تصاویر پزشکی سرطان پوست پیشنهاد می‌شود. در مدل پیشنهادی از یادگیری عمیق در دو سطح استفاده شده‌است و سپس در هر سطح از الگوریتم CatBoost برای ترکیب این مدل‌ها استفاده می‌شود. نتایج مدل پیشنهادی با شبکه‌های عمیق تک‌سطحه و پژوهش‌های مشابه پیشین مقایسه شده‌است. نتایج نشان‌می‌دهد که مدل پیشنهادی عملکرد بهتری در رده‌بندی تصاویر سرطان پوست دارد. عملکرد مدل پیشنهادی، چه در هر یک از کلاس‌ها و چه درکل، از تمامی مدل‌های یادگیری عمیق مستقل بهتر بوده‌است. همچنین نشان‌داده شده که استفاده از VGG-Ensemble در کنار روش پیشنهادی و ترکیب نتایج آن به‌کمک CatBoost و تشکیل یک مجمع دوسطحی، عملکرد آن را در هر کلاس نیز بهبود داده‌است.

واژگان کلیدی: یادگیری عمیق، یادگیری جمعی، سرطان پوست، خوش‌خیم، بدخیم.

Classification of skin cancer images using two-level ensemble deep learning

Pouyan Jaberi[1], Shahla Nemati*[2], Mohammad Ehsan Basiri3

Master Student, Department of Computer Engineering, Shahrekord University, Iran[1]
Assistant Professor, Department of Computer Engineering, Shahrekord University, Iran[2]
Associate Professor, Department of Computer Engineering, Shahrekord University, Iran[3]

Abstract

Today, despite the tremendous advances in medical science and technology, access to a specialist doctor is still considered a major challenge. This challenge is of great importance for diseases such as cancer. Skin cancer is the 13th most common cancer in men and the 15th most common cancer in women. While some skin problems are benign and harmless, some of them can be malignant masses, which will remain harmless if they are diagnosed in time. When consulting a specialist doctor may be time-consuming and expensive, an intelligent system can be a fast alternative or, at least, an efficient preliminary treatment solution. For skin cancer, such intelligent system may utilize the images of suspicious skin masses labeled according to their benign or malignant state by specialist physicians. These labeled images are useful for training intelligent systems which should diagnose the potential problems in unseen new images.

In this research, a novel deep learning-based approach is proposed for the problem of classifying skin cancer images into two categories of benign and malignant images. In the proposed model, powerful

* نویسندهٔ عهده‌دار مکاتبات

* Corresponding author

• تاریخ ارسال مقاله: ۱۴۰۱/۸/۱۹ • تاریخ پذیرش: ۱۴۰۲/۱۲/۶ • تاریخ انتشار: ۱۴۰۳/۸/۸ • نوع مطالعه: پژوهشی

[Downloaded from jsdp.rcisp.ac.ir on 2026-11-04]

[DOI: 10.61186/jsdp.21.2.79]

deep learning models for image classification including VGG, ResNet, and Inception are used in two levels. Specifically, we formed two ensembles; VGG ensemble which consists of VGG-16 and VGG-19 models and ResNet ensemble which consists of ResNet152, ResNet50, and Inception models. CatBoost algorithm is used in each level to combine the models on that ensemble. Finally, at the next level, two ensembles were combined using the CatBoost algorithm. The proposed ensemble model tries to improve the accuracy and consistency of the results by aggregating the deep models at its two levels. In order to show the utility of the proposed model, a subset of ISIC public dataset for skin cancer images is used for training and evaluation of models. The performance of the proposed ensemble model is compared with several deep neural networks and previous similar researches. Specifically, we compared the results achieved by the proposed model with those obtained by existing similar deep models and those used as building blocks of the proposed model. The results show that the proposed model performs better in classifying skin cancer images. The performance of the proposed model, both in each of the classes and in general, has been better than all independent deep learning models. It has also been shown that using VGG ensemble along with this proposed model by combining its results with the help of CatBoost and forming a two-level ensemble has improved its independent performance in each class.

Keywords: Deep Learning, Ensemble Learning, Skin Cancer, Benign, Malignant.

۱-مقدمه

برخلاف پیشرفت‌های شگرف در علوم‌پزشکی، همچنان دسترسی به پزشک و به‌ویژه پزشک متخصص نزد عموم یک چالش به‌شمار می‌رود؛ به‌ویژه که در شماری از نواحی جهان، نسبت تعداد پزشکان به جمعیت محلی بسیار کم‌تر از آستانهٔ استاندارد است[۱]. مطابق داده‌های بانک جهانی، در جهان به‌ازای هر هزار نفر جمعیت تنها ۱.۸ نفر پزشک وجود دارد که این مقدار در کشور ایران به ۱.۶ می‌رسد (این نسبت‌ها براساس آخرین داده‌های موجود تا ۲۰۲۰ محاسبه
و در شکل (۱) نشان داده شده‌است)[۱].



(شکل-۱): شمار پزشکان به هر هزار نفر جمعیت در کشورهای جهان براساس آخرین داده‌های موجود تا ۲۰۲۰ [۱].

(Figure-1): The number of doctors per thousand population in the countries of the world [۱].

تشخیص خودکار بیماری‌ها به‌صورت هوشمند می‌تواند دورهٔ طلایی درمان بیماری را زودتر تشخیص‌داده و کمک کند بیماری به‌موقع درمان شود [۲],[۳]. در سال‌های اخیر فناوری اطلاعات و به‌خصوص هوش‌مصنوعی در زمینهٔ پزشکی و مراقبت‌های بهداشتی، پیشرفت‌های زیادی برای تشخیص و پیشنهاددهی درمان داشته‌است [۴]، [۵]. معمولاً منظور از هوش‌مصنوعی در پزشکی استفاده از مدل‌های یادگیری ماشین برای جستجوی داده‌های پزشکی و کشف اطلاعاتی برای کمک به‌بهبود نتایج سلامت و تجربیات بیمار است.

در این پژوهش نمونه‌ای از زیرمسائل یادگیری ماشین یعنی رده‌بندی تصاویر سرطان پوست را بررسی خواهیم کرد؛ که یک نمونه از یادگیری نظارتی هستند. در این نوع مسائل، یک مجموعه‌داده شامل تعدادی تصویر و همچنین برچسب‌های متناظر با هر تصویر در اختیار مدل قرار می‌گیرد تا مدل بتواند با دیدن هر تصویر برچسب نظیر آن را تولید کند. ارزیابی مدل نیز با ارائهٔ مجموعهٔ تصاویری جدید که پیش‌تر توسط مدل دیده نشده‌است و مقایسهٔ برچسب‌های پیش‌بینی‌شده توسط مدل با برچسب‌های واقعی آن‌ها انجام می‌گیرد [۴], [۶]. نمای کلی مسئلهٔ رده‌بندی تصاویر سرطان پوست به‌عنوان یک مسئلهٔ یادگیری نظارتی در شکل (۲) نشان داده شده‌است.

از جملهٔ چالش‌برانگیزترین مراحل در یادگیری ماشین به‌شکل سنتی، می‌توان به مهندسی ویژگی‌ها[1] اشاره کرد که شامل شناسایی و استخراج ویژگی‌های اثرگذار در نتیجهٔ موردنظر می‌شود تا با استفاده از یادگیری ماشین بتوان میان آن‌ها و نتیجهٔ نهایی، رابطه‌ای کشف‌کرده و مدلی تشکیل‌داد، اما در رویکردهای نوین یادگیری ماشین و به‌ویژه یادگیری عمیق، تلاش براین است که این مرحله نیز به‌طور خودکار انجام شود. [۴]. یادگیری عمیق[2] تلاش می‌کند تا ویژگی‌هایی سطح بالا و عموماً انتزاعی را از داده‌های خام ورودی استخراج و سپس از آن‌ها برای انجام

[1] Feature Engineering
[2] Deep Learning

[Downloaded from jsdp.rcisp.ac.ir on 2026-11-04]

[DOI: 10.61186/jsdp.21.2.79]

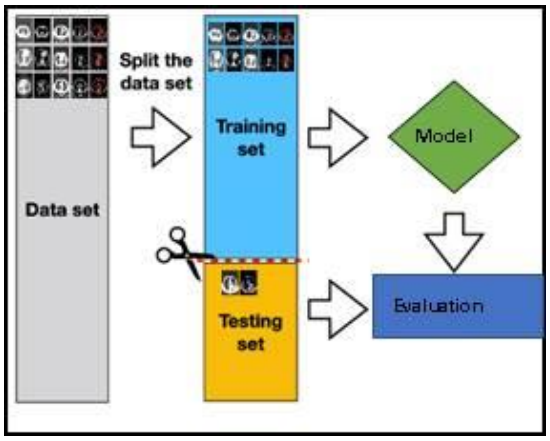
عملیاتی چون خوشه‌بندی، رده‌بندی و... استفاده کند[۷]، [۸].

درخصوص تصاویر، خودکارسازی مرحلهٔ مهندسی ویژگی‌ها به‌این معناست که مدل بتواند تشخیص دهد که به‌کدام جزئیات موجود در تصویر ورودی توجه‌کند و چه چیزهایی را نادیده بگیرد، اما مشکل اصلی این است که اجرای چنین عملیاتی با شبکه‌های عصبی تمام‌متصل[1] بسیار هزینه‌بر انجام می‌پذیرد؛ چراکه:

۱-تک‌تک پیکسل‌های موجود در تصویر باید به‌عنوان ورودی به شبکه راه پیدا کنند و برای تصاویر رنگی نیز هر کانال رنگ، خود نیازمند ورودی مجزا است و باتوجه به‌اینکه دریک شبکهٔ تمام‌متصل، هر نورون به‌تمامی نورون‌های لایهٔ پس از خود متصل است، تعداد اتصالات و در نتیجه تعداد پارامترها، بسیار زیاد می‌شود.

۲-در ورودی یک شبکهٔ تمام‌متصل، تصاویر دوبُعدی (و چندبُعدی – مثل تصاویر MRI) به ماتریس‌هایی یک‌بُعدی تبدیل می‌شوند؛ لذا شبکه نمی‌تواند از تجربه‌ای که در یافتن ویژگی‌ها در قسمتی خاصی از یک تصویر به‌دست آورده‌است، برای یافتن همان ویژگی‌ها در قسمت دیگری از یک تصویر دیگر نیز استفاده کند.

نتیجه آنکه ضروری است به‌شکلی بتوان ابعاد تصاویر را کاهش داد، ویژگی‌های اصلی موجود در تصاویر را نیز حفظ و یا حتّی به‌شکلی رمزنگاری کرد [۹].



(شکل-۲): نمای کلی مسئلهٔ رده‌بندی تصاویر سرطان پوست به‌عنوان یک مسئلهٔ یادگیری نظارتی [۶].

(Figure-2): Overview of the skin cancer classification problem as a supervised learning problem [6] .

یکی از بهترین روش‌های استخراج ویژگی که به‌ویژه در داده‌های تصویری نیز کاربرد گسترده‌ای دارد، شبکه‌های عصبی کانولوشن[2] (CNN) است[۱۰]. در این نوع شبکه، استخراج ویژگی‌ها از طریق قراردادن پالایه‌هایی بر روی داده‌های ورودی انجام‌شده‌است و به‌طورمعمول با استفاده از لایه‌های ادغام نیز کاهش ابعاد انجام می‌پذیرد. درنهایت، آنچه حاصل می‌شود یک نقشهٔ ویژگی بوده که دارای ابعادی بسیار کوچک‌تر از دادهٔ ورودی و دربردارندهٔ ویژگی‌های اصلی موجود در داده‌است. حال بسته به‌هدف، می‌توان این نقشهٔ ویژگی را به‌عنوان ورودی برای یک شبکهٔ عصبی دیگر درنظرگرفت.

مشکل بعدی در شبکه‌های عصبی عمیق بزرگ استفاده‌شده در پردازش تصاویر دقت و اطمینان شبکه است. برای پاسخ به این مشکل از رویکردهای متفاوتی در یادگیری ماشین استفاده شده‌است. یکی از این رویکردها یادگیری جمعی[3] است که تلاش می‌کند چندین مدل یادگیرنده را هم‌زمان به‌کارگرفته و نتیجهٔ نهایی را از به‌هم پیونددادن نتایج مدل‌های جداگانه به‌دست‌آورد [۱۱].

یادگیری جمعی عمیق روشی برای بهره‌مندی هم‌زمان از مزایای یادگیری جمعی و یادگیری عمیق است؛ به‌این‌سان که چند مدل یادگیری عمیق در کنار یکدیگر به‌کار گرفته‌شده و نتیجهٔ نهایی با استفاده از روش‌های مرسوم در یادگیری جمعی فرآوری می‌شود. ایدهٔ بنیادین این است که یک مدل استوار بر یادگیری جمعی عمیق، در سنجش با مدل‌های مستقل، تعمیم‌پذیری بیش‌تر و برایندی بهتر را به‌همراه داشته‌باشد[۱۲].

در پژوهش حاضر، باتوجه به‌قدرت بالای مدل‌های مبتنی بر شبکه‌های عصبی کانولوشن برای استخراج ویژگی، مدلی ترکیبی برای به‌کارگیری پنج نوع شبکهٔ کانولوشنی برتر پیشنهاد خواهد شد؛ همچنین برای استفادهٔ هم‌زمان مزایای یادگیری جمعی و یادگیری عمیق از یادگیری جمعی عمیق استفاده خواهد شد.

در این پژوهش، روشی برای پیکربندی و پیاده‌سازی یک مدل مبتنی بر یادگیری جمعی عمیق برای رده‌بندی تصاویر پزشکی سرطان پوست پیشنهاد شده‌است. در مدل پیشنهادی یادگیری عمیق ترکیبی دوسطحی براساس پنج مدل کانولوشنی رایج و برتر سال‌های اخیر پیشنهاد شده‌است. به‌صورت دقیق‌تر مدل‌ها را به‌دو مجمع مجزا تقسیم‌کرده که خودِ مجامع دریک سطح بالاتر تشکیل مجمع جدیدی می‌دهند؛ یعنی خروجی مجامع سطح نخست خود به‌عنوان ورودی یک الگوریتم ترکیب‌کنندهٔ

[1] Fully Connected
[2] Convolutional Neural Network (CNN)
[3] Ensemble Learning

[Downloaded from jsdp.rcisp.ac.ir on 2026-11-04]
[DOI: 10.61186/jsdp.21.2.79]

دیگر درنظر گرفته‌شده و به‌این‌ترتیب مجمعی دولایه داریم. همچنین برای ترکیب‌کنندهٔ هر مجمع نیز برخلاف پژوهش‌های پیشین دوباره از یادگیری ماشین استفاده کردیم. انگیزهٔ استفاده از پنج نوع شبک کانولوشنی برتر، قدرت بالای مدل‌های مبتنی بر شبکه‌های عصبی کانولوشن برای استخراج ویژگی است. همچنین انگیزهٔ استفاده از یادگیری جمعی عمیق نوآوری پژوهش حاضر نسبت به پژوهش‌های پیشین موارد زیر است:

- برخلاف پژوهش‌های پیشین و همچنین با هدف افزایش بایاس، در مدل پیشنهادی در سطح نخست مجمع، مدل‌های هم‌خانواده با معماری‌های مشابه دریک گروه قرار می‌گیرند.
- خروجی ترکیب‌کننده‌های مجامع VGG و ResNet خود به‌عنوان ورودی یک الگوریتم ترکیب‌کنندهٔ دیگر درنظر گرفته‌شده و به‌این ترتیب مجمعی دولایه پیشنهاد شده‌است.
- همچنین در روش پیشنهادی، برای ترکیب‌کنندهٔ هر مجمع نیز برخلاف دو پژوهش پیشین و استفاده از روشی چون رأی اکثریت یا میانگین‌گیری وزن‌دار، دوباره از یادگیری ماشین استفاده‌شده و الگوریتم CatBoost انتخاب شده‌است.

۲-پژوهش‌های پیشین

به طور کلی دو دسته از پژوهش‌های پیشین مورد بررسی قرار گرفتند:

- پژوهش‌هایی درخصوص عملکرد یادگیری عمیق جمعی (به‌ویژه در رده‌بندی تصاویر پزشکی)؛
- پژوهش‌هایی که از شبکه‌های عصبی عمیق مطرح برای رده‌بندی تصاویر پزشکی استفاده کرده‌اند.

بدیهی است که پژوهش‌های پرشماری در این دو مورد می‌توان یافت؛ لذا دراین پژوهش، اغلب بر پژوهش‌های سال‌های اخیر و همچنین مقالات جامعی که شمار زیادی از پژوهش‌های مرتبط را بررسی کرده‌اند، تمرکز شده‌است که در ادامه به اختصار به آن‌ها اشاره خواهد شد.

۲-۱-یادگیری ژرف جمعی

پژوهش گانا و همکاران [۱۲]، از معدود پژوهش‌هایی است که به‌مرور جامع یادگیری عمیق جمعی و بررسی و مقایسهٔ شاخص‌ترین مدل‌های مبتنی بر این روش پرداخته‌است. دراین مقاله، پس از بررسی مختصر نظریهٔ یادگیری جمعی، روش‌های گوناگونی جهت تشکیل یک مجمع از جمله دسته‌بندی، توان‌افزایی و انباشتگی را بازبینی‌کرده و همچنین استفاده از مدل‌های یادگیری عمیق در مجمع‌ها را مورد بررسی قرار داده‌است.

کینی و همکاران [۱۳] اقدام به‌توسعهٔ یک چارچوب به‌منظور تشخیص خودکار ابتلا به بیماری کووید-۱۹ برپایهٔ تصاویر CT سینه کرده‌اند که در آن تجهیزات پزشکی مجهز به اینترنت اشیا[1] (IoT) متصل به شبکه‌ای بوده که از طریق آن داده‌ها خود را به‌یک لایهٔ ذخیره‌سازی داده ارسال می‌کنند. تشخیص بیماری در لایهٔ مجزایی موسوم به لایهٔ پردازش‌داده انجام‌شده و نتایج آن دریک لایهٔ کاربردی به‌دست کاربران (ازجمله پزشکان) می‌رسد.

لایهٔ پردازش‌داده در این پژوهش از یک مجمع یادگیری عمیق تک‌سطحی تشکیل‌شده‌است، که از سه مدل ResNet152V2، DenseNet201 و InceptionResNetV2 بهره می‌برد؛ همچنین ترکیب نتایج سه مدل از روش رأی نسبی اکثریت انجام می‌شود. به‌بیانی دیگر، در ابتدا سه مدل مبتنی بر شبکه‌های عصبی کانولوشنی تصاویر را به‌عنوان ورودی دریافت‌کرده و یک پیش‌بینی دربارهٔ ردهٔ هریک انجام می‌دهند) برای مثال اینکه آیا تصویر CT موردنظر متعلق به بیماری مبتلا به کووید-۱۹ است یا خیر .(تصمیم نهایی براساس رأی نسبی اکثریت میان این سه مدل گرفته می‌شود که این بخش، ترکیب‌کننده نام دارد[۱۳].

در پژوهشی مشابه، داس و همکاران [۱۴] برای رده‌بندی تصاویر X-ray سینه برای تشخیص ابتلا به کووید-۱۹ از سه مدلDenseNet201 ، ResNet50V2و InceptionV3 استفاده کرده‌اند. این‌بار نتیجهٔ نهایی با نوعی میانگین‌گیری وزن‌دار انجام می‌شود که توسط همین پژوهش معرفی شده‌است. گفتنی است که دو مدل معرفی‌شده در [۱۳], [۱۴] به‌دلیل تشابه با مدل پیشنهادی در پژوهش جاری، گزینه‌هایی برای مقایسه با روش پیشنهادی درنظر گرفته‌شده‌اند.

۲-۲-تشخیص تصاویر پزشکی با یادگیری عمیق

در [۱۵] مرور جامعی بر پژوهش‌های سال‌های اخیر در زمینهٔ تشخیص سرطان پوست با روش‌های یادگیری عمیق انجام شده‌است. این بررسی نشان‌دهندهٔ نقش بسیار پررنگ شبکه‌های عصبی کانولوشنی در بررسی و به‌ویژه رده‌بندی تصاویر پزشکی سرطان پوست با کمک یادگیری عمیق است.

[1] Internet of Things

[Downloaded from jsdp.rcisp.ac.ir on 2026-11-04]

[DOI: 10.61186/jsdp.21.2.79]

در [۱۶] کاسم و همکاران، با به‌کارگیری شبکهٔ GoogLeNet برای استخراج ویژگی و یک مدل SVM برای رده‌بندی، اقدام به‌توسعهٔ مدلی برای رده‌بندی تصاویر سرطان پوست در هشت کلاس از مجموعه‌دادهٔ ISIC 2019 کرده‌است. نویسندگان در [۱۷]، یک شبکهٔ ۱۶ VGG-را بر روی ۱۲۸۰ تصویر سرطان پوست از آرشیو ISIC ، برای رده‌بندی در دو کلاس، آموزش داده و ارزیابی کرده‌اند.

همچنین در [۱۸] عملکرد شبکه‌های ResNet50، InceptionV3، DenseNet169، InceptionResNetV2 و MobileNet را بر روی ۳۶۰۰ تصویر از آرشیو ISIC در یک کلاس (بدخیم یا غیربدخیم) مقایسه کرده‌است. بیشترین دقت با استفاده از ResNet50 حاصل شده‌است. مختصری از این پژوهش‌ها در جدول (۱) آورده شده‌است. همان‌گونه که در مقدمه ذکر شد، تمرکز پژوهش جاری بر روی معرفی روشی نوین برمبنای یادگیری عمیق جمعی است؛ به‌منظور آزمایش و ارزیابی این روش، تصاویر سرطان پوست استفاده شده‌اند.

(جدول-۱): مدل‌های مورد بررسی در پژوهش‌های پیشین.
(Table-1): The models examined in previous researches.

پژوهش	توضیحات	مجموعهٔ داده‌ها	سال
[۱۳]	**ResNet152V2** ، **DenseNet201** و **Inception-ResNetV2** با استفادهٔ رأی نسبی اکثریت	تصاویر **CT** سینه	۲۰۲۲
[۱۴]	**DenseNet201** ، **ResNet50V2** و **InceptionV3** با میانگین‌گیری وزن‌دار	تصاویر **X-ray** سینه	۲۰۲۱
[۱۲]	مقالهٔ مروری (روش‌های یادگیری عمیق جمعی)	-	۲۰۲۱
[۱۵]	مقالهٔ مروری (تشخیص سرطان پوست با یادگیری عمیق)	-	۲۰۲۱
[۱۶]	**GoogLeNet+SVM**	**ISIC 2019**	۲۰۲۰
[۱۷]	**VGG-16**	آرشیو **ISIC**	۲۰۱۶
[۱۸]	**CNN** و چند **ResNet50** دیگر برای مقایسه	آرشیو **ISIC**	۲۰۲۱

۳-مجموعه‌داده

مجموعه‌دادهٔ اصلی مورد استفاده، زیرمجموعه‌ای از (آرشیو ISIC) مرتبط با تصاویر پزشکی سرطان پوست است که درمجموع شامل ۳۲۹۷ تصویر رنگی (سه‌کاناله) ۲۲۴×۲۲۴ پیکسل می‌شود. این تصاویر به دو ردهٔ خوش‌خیم و بدخیم تقسیم می‌شوند (به‌ترتیب شامل ۱۸۰۰ و ۱۴۹۷ تصویر).

نمونه‌ای از تصاویر ردهٔ خوش‌خیم و بدخیم در شکل‌های (۳) و (۴) نشان داده شده‌است. این مجموعه‌داده به‌صورت آزاد در [۱۹] در دسترس است.

درپژوهش حاضر، ۴۰٪ مجموعه‌داده (۹۲۶ داده از ردهٔ خوش‌خیم و ۶۰۵ داده در ردهٔ بدخیم) برای آموزش و۶۰٪ باقی‌مانده (۱۱۰۴ داده از ردهٔ خوش‌خیم و ۸۷۴ داده در ردهٔ بدخیم، یا به‌ترتیب۵۶٪ و ۴۴٪) که هرگز در طی فرایند یادگیری به مدل‌ها نشان داده‌نشده، به‌منظور ارزیابی استفاده شده‌است. از ۴۰٪ داده‌های آموزشی مدل‌های پایه برروی ۲۰٪ تصاویر برچسب‌دار موجود (۶۶۰ تصویر، شامل ۳۳۰ تصویر از هر رده، که تصادفی از مجموعه‌دادهٔ اصلی انتخاب شدند) آموزش دیده‌اند و ۲۰٪ باقی‌مانده از داده‌های آموزشی برای آموزش مدل‌های CatBoost استفاده شده‌اند.

۴-معماری مدل پیشنهادی

در این پژوهش با بهره‌گیری از پنج مدل کانولوشنی رایج و برتر سال‌های اخیر معماری پیشنهادی یادگیری عمیق ترکیبی دوسطحی که در **Error! Reference source not found.** نشان داده شده‌است پیشنهاد می‌شود. مدل‌های کانولوشنی استفاده‌شده در معماری پیشنهادی در ادامه معرفی خواهند شد.

۴-۱- مدل‌های پایهٔ استفاده‌شده در معماری پیشنهادی

- **VGG-16**

این شبکه در رقابت ImageNet سال ۲۰۱۴ جزو پنج شبکهٔ برتر قرارگرفت. ورودی این شبکه، تصاویر رنگی (سه کانالهٔ RGB) در ابعاد ۲۲۴×۲۲۴ بوده و همچنین این شبکه به‌طور پیش‌فرض از ۱۶ لایه تشکیل شده‌است[۱۷].

- **VGG-19**

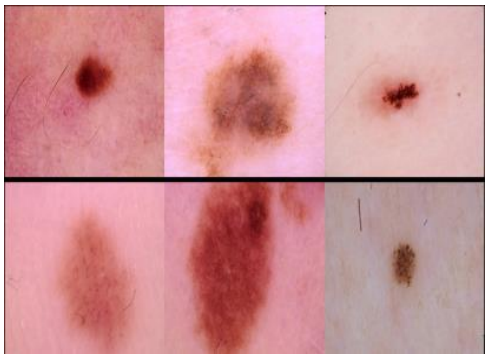
دیگر شبکهٔ برتر در رقابت ImageNet سال ۲۰۱۴، شبکهٔ VGG-19 بود که در واقع گونهٔ ارتقا یافته‌ای ازVGG-16 بوده که شمار لایه‌های کانولوشن آن به ۱۹عدد افزایش یافته‌است [۲۰].

- **ResNet-50V2**

شبکهٔ ResNet50 از جمله شبکه‌های شناخته شده‌است و در رقابت ImageNet سال ۲۰۱۵ نیز از جمله شبکه‌های برتر بود. ResNet50V2 گونهٔ اصلاح‌شده‌ای از

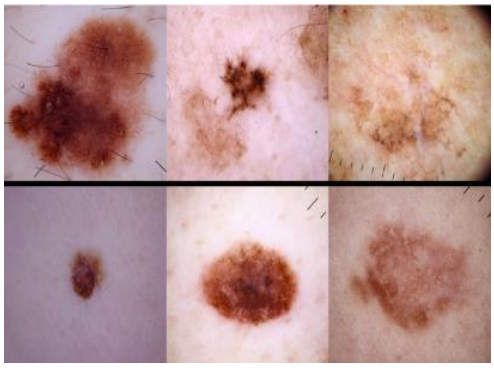
[Downloaded from jsdp.rcisp.ac.ir on 2026-11-04]
[DOI: 10.61186/jsdp.21.2.79]

این شبکه است که عملکرد بهتری در قیاس با ResNet50 و ResNet101 دارد. ورودی این شبکه، ۲۲۴×۲۲۴×۳ واحد است. همچنین ایدهٔ اصلی خانوادهٔ ResNet استفاده از روشی موسوم به اتصال جهشی[1] بوده که در آن خروجی برخی از لایه‌ها علاوه‌بر لایهٔ بعدی خود، جهشی به‌طول چندلایه نیز داشته و به‌عنوان ورودی یک لایهٔ ژرف‌تر نیز درنظر گرفته می‌شود. این امر در سهولت بهینه‌سازی شبکه و افزایش دقت (در قیاس با شبکه‌ای مشابه، اما بدون استفاده از این روش) مؤثر است [۱۸].



(شکل-۱): نمونه‌هایی از تصاویر ردهٔ خوش‌خیم.

(Figure-3): Examples of benign category images.



(شکل- ۲): نمونه‌هایی از تصاویر ردهٔ بدخیم.

(Figure-4): Examples of malignant category images.

- **ResNet152V2**

شبکهٔ ResNet152 موفق به‌کسب مقام نخست در رقابت ImageNet سال ۲۰۱۸ در زمینهٔ دفاع خصمانه شده‌است. معماری آن شامل ۱۵۲ لایه بود که یکی از ژرف‌ترین شبکه‌های مورد استفاده در این پژوهش محسوب می‌شود. این شبکه ۱۵۲ لایهٔ کانولوشن دارد که به‌صورت یک‌درمیان اتصال جهشی انجام می‌دهند [۲۱].

- **InceptionV3**

این شبکه در دسامبر ۲۰۱۵ در مقام نخست رقابت ImageNet قرارگرفت و توسعهٔ آن توسط شرکت گوگل انجام شده‌است. Inception V3 درجمع از یازده ماژول تشکیل‌شده که هریک دربرگیرندهٔ لایه‌های ادغام و پالایه‌های کانولوشن با تابع ReLU است[۲۲].

۲-۴- ترکیب مدل‌ها در معماری پیشنهادی

ضمن استفاده از روش یادگیری انتقالی [۲۳]، مدل‌های پایه بر روی ٪۲۰ تصاویر برچسب‌دار موجود (۶۶۰تصویر، شامل ۳۳۰تصویر از هر رده، که به‌صورت تصادفی از مجموعه‌دادهٔ اصلی انتخاب شدند) آموزش دیده‌اند.

برخلاف پژوهش‌های پیشین و همچنین با هدف افزایش بایاس، پیکربندی سطح نخست مجمع به‌این شکل انجام می‌شود که مدل‌های هم‌خانواده با معماری‌های مشابه در کنار یکدیگر و دریک گروه قرار می‌گیرند؛ لذا به‌جای اینکه خروجی نهایی براساس رأی اکثریت یا میانگین‌گیری وزن‌دار از تمامی مدل‌های یادشده به‌دست آید، این مدل‌ها در مدل پیشنهادی به‌دو مجمع مجزا به‌شکل زیر تقسیم شده‌است:

- VGG-Ensemble:
- VGG-16
- VGG-19
- ResNet-Ensemble:
- ResNet50V2
- ResNet152V2
- InceptionV3

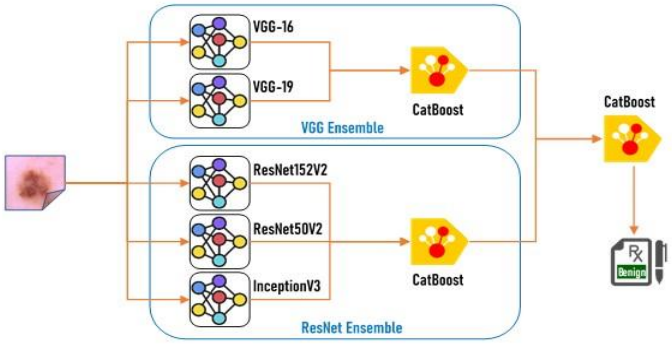
سپس خودِ مجامع ResNet و VGG در یک سطح بالاتر تشکیل مجمع جدیدی می‌دهند؛ یعنی خروجی ترکیب‌کننده‌های مجامع VGG و ResNet خود به‌عنوان ورودی یک الگوریتم ترکیب‌کنندهٔ دیگر درنظر گرفته‌شده و به این ترتیب مجمعی دولایه داریم.

همچنین برای ترکیب‌کنندهٔ هر مجمع نیز برخلاف دو پژوهش پیشین و استفاده از روشی چون رأی اکثریت یا میانگین‌گیری وزن‌دار، مجدداً از یادگیری ماشین استفاده کرده و الگوریتم CatBoost انتخاب شده‌است. این مدل یادگیری جمعی توسط شرکت یاندِکْس توسعه داده‌شده و از روش تقویت گرادیان بهره می‌برد [۲۳]. از دلایل مهم انتخاب این مدل در پژوهش جاری، عملکرد به‌نسبه بهتر آن در برابر مدل‌های مشابه، به‌ویژه هنگام کار با مجموعه‌های دادهٔ کوچک‌تر است [۲۳]. همچنین در سنجش تجربی انجام‌شده در [۲۴]، عملکرد CatBoost در برابر دو الگوریتم محبوب مشابه متن‌باز دیگر در همین حوزه XGBoost و LightGBM مورد سنجش قرارگرفته و نتیجهٔ بهتری نشان داده‌است. این‌بار نیز ٪۲۰ دیگری از داده‌ها برای آموزش مدل‌های CatBoost استفاده شده‌اند.

[1] Skip Connection

[Downloaded from jsdp.rcisp.ac.ir on 2026-11-04]

[DOI: 10.61186/jsdp.21.2.79]



(شکل-۵): معماری مدل پیشنهادی.

(Figure-5): Architecture of the proposed

۵- روش مقایسه

برای بررسی و ارزیابی عملکرد مدل پیشنهادی حالت‌های زیر درنظر گرفته شده‌است:

۱- هفت مدل CNN مستقل بر روی مجموعه‌داده‌های همین پژوهش آموزش داده می‌شوند تا اساساً عملکرد مجمع‌ها در برابر عملکرد مدل‌های مستقل سنجیده شود. این مدل‌ها شامل تمامی پنج CNN مورد استفاده در سطح نخست مدل پیشنهادی (VGG-16، VGG-19، ResNet-50V2، ResNet-152V2 و InceptionV3) به‌علاوهٔ DenseNet201 و InceptionResNetV2 می‌شوند.

۲- یک مجمع تک‌سطحی از همان پنج CNN مدل پیشنهادی تشکیل‌شده و خروجی هر CNN به‌عنوان ورودی یک مدل CatBoost درنظر گرفته‌می‌شود که این مدل در شکل (۶) نشان داده شده‌است.

۳- مدل‌های ارائه‌شده در پژوهش‌های [۱۳]، [۱۴] از نو بازسازی‌شده و بر روی مجموعهٔ داده‌های همین پژوهش آموزش می‌بینند.

پارامتر مدل‌های CNN براساس [۲۵] و به‌صورت زیر انتخاب شده‌است:

- تابع اتلاف[1]: به‌دلیل وجود دو کلاس در مسئله، از تابع Binary Crossentropy استفاده شده‌است.
- بهینه‌ساز[2]: از الگوریتم آدام[3] استفاده شده‌است.
- توابع فعال‌ساز[4] لایه‌های هر لایهٔ CNN: بدون هیچ تغییری و دقیقاً مطابق پیاده‌سازی پیش‌فرض هر مدل تعیین شده‌است.
- تابع فعال‌ساز لایهٔ تمام‌متصل خروجی: به‌دلیل وجود دو کلاس در مسئله، تابع سیگموئید[5] به‌کار رفته‌است.

[1] Loss Function
[2] Optimizer
[3] Adam
[4] Activation Function
[5] Sigmoid

به طور معمول در آمار جهت سنجش کارایی یک مدل در رده‌بندی دودویی دو معیار دقت[6] (P) و بازیابی[7] (R) به‌کار می‌روند که به‌ترتیب به‌صورت روابط (۱) و (۲) نشان داده می‌شوند [۲۶]:

$$\boldsymbol{P} = \frac{\text{مثبت صحیح}}{\text{مثبت کاذب+مثبت صحیح}} \qquad (۱)$$

$$\boldsymbol{R} = \frac{\text{مثبت صحیح}}{\text{منفی کاذب+مثبت صحیح}} \qquad (۲)$$

به بهای چشم‌پوشی از بازیابی، می‌توان دقت را افزایش داد (و برعکس)؛ لذا توجه انحصاری به‌هریک از دو معیار می‌تواند نتایجی گمراه‌کننده داشته باشد؛ پس نیاز است معیاری به‌کار ببریم که بتواند هر دومعیار یادشده را ترکیب‌کند و به‌عنوان شاخصی برای سنجش صحت به‌کار رود. برای این امر از امتیاز F1[8] استفاده می‌کنیم که در رابطهٔ (۳) نشان داده شده‌است:

$$\boldsymbol{F_1} = \frac{\boldsymbol{2.P.R}}{\boldsymbol{P+R}} \qquad (۳)$$

امتیاز F1 نوعی میانگین هارمونیک بوده که به‌طور معمول و همچنین در بیشتر مقالات بررسی‌شده دراین پژوهش، برای سنجش صحت رده‌بندهای دودویی مورداستفاده قرارگرفته؛ لذا در این پژوهش نیز این معیار برای بررسی نهایی عملکرد مدل‌ها به‌کار خواهد رفت.

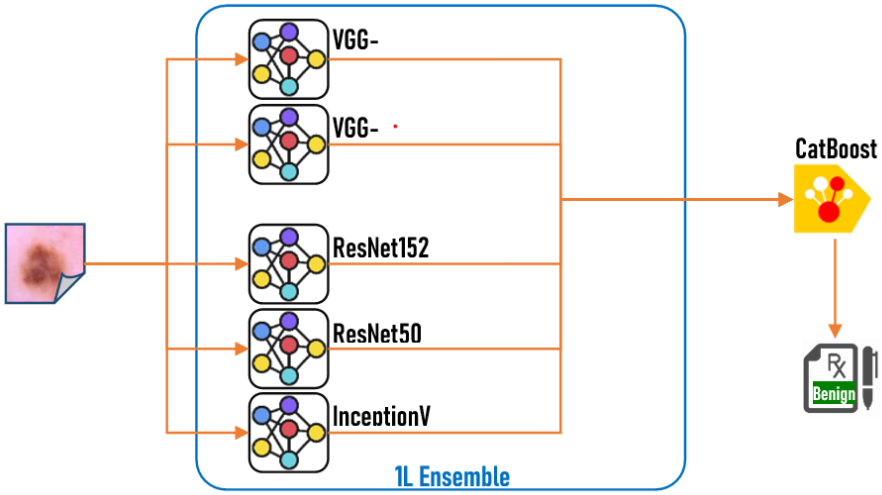
۶- نتایج ارزیابی

از مجموع تمام ۳۲۹۸ داده‌های برچسب‌دار تصویری ۶۰٪ باقی‌مانده (۱۱۰۴ داده از ردهٔ خوش‌خیم و ۸۷۴ داده در ردهٔ بدخیم، یا به‌ترتیب۵۶٪ و ۴۴٪) که هرگز در طی فرایند یادگیری به مدل‌ها نشان داده‌نشده، به‌منظور ارزیابی استفاده شده‌است. امتیاز F1 برای هر مدل و درهر کلاس محاسبه و در **Error! Reference source not found.** و شکل (۷) نمایش داده شده‌است.

[6] Precision
[7] Recall
[8] F1-score

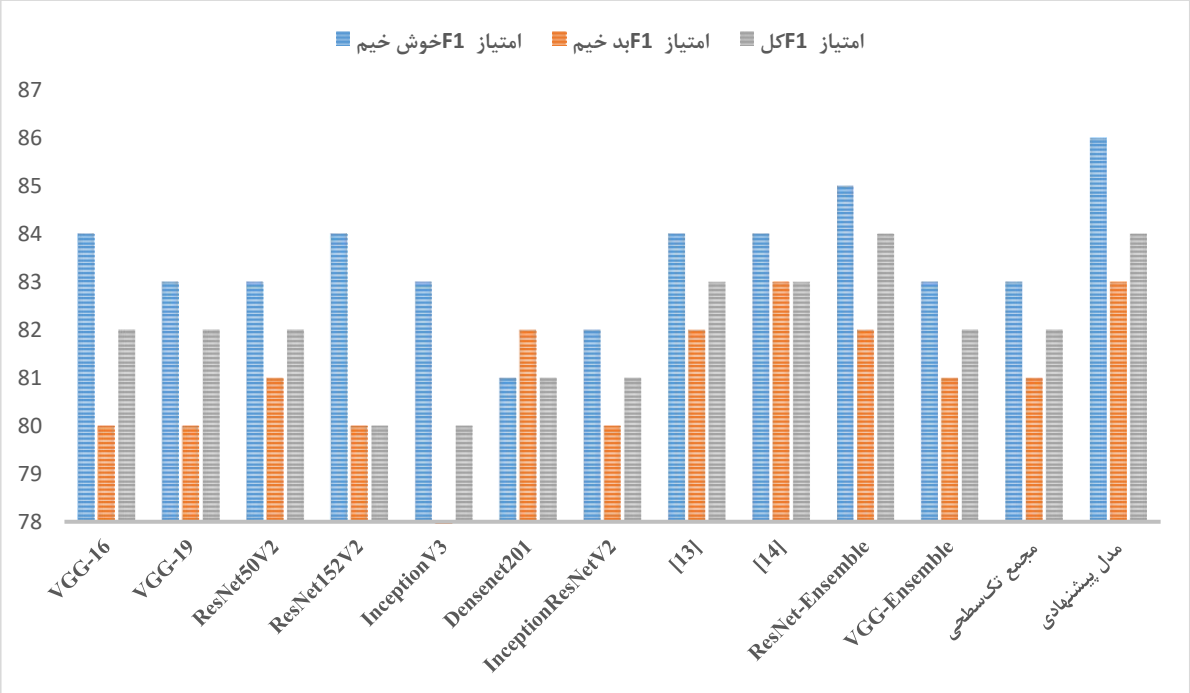
[Downloaded from jsdp.rcisp.ac.ir on 2026-11-04]

[DOI: 10.61186/jsdp.21.2.79]



(شکل-۶): معماری مجمع تک‌سطحی.

(Figure-6): Single-level ensemble architecture.



(شکل-۷): شکل نتایج نهایی.

(Figure-7): Illustration of the final results.

از نتایج، نخست می‌توان دریافت که استفاده از روش یادگیری جمعی به‌طورمعمول نتیجهٔ بهتری از به‌کارگیری مدل‌های مستقل دارد. امتیاز F1 کلی هیچ‌یک از مدل‌های مستقل بهتر از مجامع نبود. تنها شماری از مدل‌های مستقل را می‌توان یافت که به‌صورت موردی و در یک کلاس، عملکرد بهتری از برخی از مجامع داشته باشند، برای مثال: پیش‌بینی‌های مدل VGG-16 در کلاس خوش‌خیم، دقیق‌تر از پیش‌بینی‌های VG-Ensemble بودند (درمقابل اما VGG-Ensemble در کلاس بدخیم عملکرد بهتری از هر دو مدل تشکیل‌دهندهٔ خود، VGG-16 و VGG-19، نشان‌داد)؛ درحالی‌که، ازسوی دیگر برای نمونه می‌توان ResNet-Ensemble را مثال زد که هم دقت کلی آن و هم دقت آن در هر کلاس، از هردو مدل تشکیل‌دهندهٔ آن؛ یعنی ResNet50V2 و ResNet152V2 بیشتر است.

[Downloaded from jsdp.rcisp.ac.ir on 2026-11-04]

[DOI: 10.61186/jsdp.21.2.79]

(جدول-۲): نتایج نهایی براساس امتیاز F1.
(Table-2): Final results based on F1-score

نام مدل	F1امتیاز خوش‌خیم	F1امتیاز بدخیم	کل F1امتیا	توضیحات
VGG-16	۸۴	۸۰	۸۲	
VGG-19	۸۳	۸۰	۸۲	
ResNet50V2	۸۳	۸۱	۸۲	
ResNet152V2	۸۴	۸۰	۸۰	
InceptionV3	۸۳	۷۶	۸۰	
Densenet201	۸۱	۸۲	۸۱	
InceptionResNetV2	۸۲	۸۰	۸۱	
[۱۳]	۸۴	۸۲	۸۳	
[۱۴]	۸۴	۸۳	۸۳	بدخیم به همراه مدل پیشنهادی F1بالاترین امتیاز
ResNet-Ensemble	۸۵	۸۲	۸۴	کل، به همراه مدل پیشنهادی F1بالاترین امتیاز
VGG-Ensemble	۸۳	۸۱	۸۲	
مجمع تک‌سطحی	۸۳	۸۱	۸۲	
مدل پیشنهادی	**۸۶**	**۸۳**	**۸۴**	**در خوش‌خیم، بدخیم و کل F1بالاترین امتیاز**

این مثال خوبی است که چگونه دو مدل ضعیف‌تر می‌توانند خطاهای یکدیگر را جبران کرده و در قالب یک مجمع به‌نتایج بهتری دست پیدا کنند؛ لذا همچنان می‌توان از عملکرد یادگیری عمیق جمعی دفاع کرد و یا کمینه عملکرد آن را در برابر مدل‌های مستقل مورد آزمون قرارداد (با این فرض که به‌احتمال نتایج بهتری برای ارائه خواهد داشت) و اما درخصوص مدل پیشنهادی، از نتایج می‌توان دید که:

۱- عملکرد آن، چه درهر یک از کلاس‌ها و چه درکل، از تمامی CNNهای مستقل بهتر بوده‌است.

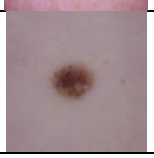
۲- با آنکه عملکرد کلی آن برابر ResNet-Ensemble است، اما استفاده از VGG-Ensemble در کنار این مجمع و ترکیب نتایج آن به‌کمک CatBoost و تشکیل یک مجمع دوسطحی، عملکرد مستقل آن در هر کلاس را نیز بهبود داده‌است.

۳- عملکرد آن، بازهم چه درهر یک از کلاس‌ها و چه درکل، از عملکرد مجمع همتای تک‌سطحی آن بهتر است.

۷- مطالعهٔ موردی

برای اینکه تحلیل جامع‌تری از تفاوت تشخیص مدل پیشنهادی دوسطحی و دو مدل معرفی‌شده در [۱۳], [۱۴] داشته باشیم، یک مطالعهٔ موردی ارائه می‌شود که در آن پنج نمونه تصویر مورد بررسی قرارگرفتند. نتایج در جدول (۳) نشان داده شده‌است.

(جدول-۳): مطالعهٔ موردی. پنج نمونه تصویر از نمونه‌های مثبت (بدخیم) و منفی (خوش‌خیم) که نتیجهٔ مدل پیشنهادی و مدل‌های [۱۳], [۱۴] متفاوت بوده‌است.
(Table-3): Case study. Five positive (malignant) and negative (benign) image samples from dataset on which the output of our proposed model and the two models presented in [13] and [14] are different.

تصویر نمونه	واقعیت	مدل پیشنهادی	[۱۳]	[۱۴]
	منفی	منفی	مثبت	مثبت
	منفی	منفی	مثبت	مثبت
	منفی	منفی	مثبت	مثبت
	مثبت	مثبت	منفی	منفی
	مثبت	مثبت	منفی	منفی

[Downloaded from jsdp.rcisp.ac.ir on 2026-11-04]
[DOI: 10.61186/jsdp.21.2.79]

همان‌گونه که در جدول (۳) مشخص شده، نتایج مدل پیشنهادی در موارد نشان داده شده به‌طورکامل متفاوت با مدل‌های مقایسه شده‌است. همچنین تشخیص درست در این موارد با مدل پیشنهادی بوده‌است.

از آنجا که ساختار درونی لایه‌های این سه مدل شبیه است، تفاوت پاسخ‌ها را می‌توان در دوسطحی‌بودن تصمیم‌گیری مدل پیشنهادی و استفاده از روش تجمیع مبتنی بر یادگیری ماشین در مدل پیشنهادی نسبت به روش‌های استفاده‌شده در [۱۳]، [۱۴] دانست.

۷-۱- پیچیدگی محاسباتی مدل پیشنهادی

در مقایسه با الگوریتم‌های مقایسه‌شده که به‌صورت تکی استفاده شده‌اند؛ یعنی VGG-16، VGG-19، ResNet-50V2، ResNet-152V2 و InceptionV3 قطعاً پیچیدگی محاسباتی مدل پیشنهادی و مدل‌های مجمع تک‌سطحی بالاتر است. دلیل این امر استفادهٔ سری (متوالی) پنج مورد از مدل‌ها در مجامع است. به‌صورت خاص اگر پیچیدگی محاسباتی مدل‌های VGG-16، VGG-19، ResNet-50V2، InceptionV3 و ResNet-152V2 را به‌ترتیب m_1 تا m_5 بنامیم و پیچیدگی مدل‌های DenseNet201 و InceptionResNetV2 را نیز m6 و m7 بنامیم، پیچیدگی مدل تک‌سطحی پیشنهادی به‌صورت رابطهٔ (۴) نشان داده شده‌است:

$$O(CatBoost) + \sum_{i=1}^{5} O(m_i) \qquad (۴)$$

همچنین پیچیدگی مدل‌های ارائه‌شده در مراجع [۱۳] و [۱۴] به‌صورت زیر است:

$$\sum_{m\in\{m_3,m_4,m_6\}} O(m) \text{ و } \sum_{i=4}^{7} O(m_i) \qquad (۵)$$

که البته یک O(n) به‌خاطر میانگین‌گیری نهایی یا استفاده از رأی‌گیری در لایهٔ آخر مدل‌ها باید با مقادیر فوق جمع شود که به‌دلیل غلبه پیچیدگی مدل‌های m_1 تا m_7 بر O(n) آن را حذف کردیم. در مورد مدل دوسطحی پیشنهادی چون سه‌بار (یک‌بار در هر مجمع و یک برای در سطح دوم) از الگوریتم CatBoost استفاده شده پیچیدگی محاسباتی کلی به‌صورت زیر است:

$$3 \times O(CatBoost) + \sum_{i=1}^{5} O(m_i) \qquad (۶)$$

که درنهایت باتوجه به‌اینکه ضریب در پیچیدگی نهایی تأثیرگذار نیست می‌توان گفت همان مقدار مربوط به مجمع تک‌سطحی یعنی همان مقدار اشاره‌شده در رابطهٔ (۴) خواهد بود.

در مورد الگوریتم CatBoost نیز همان‌گونه که در [۲۴] اشاره شده‌است پیچیدگی بخش‌های مختلف الگوریتم شامل محاسبهٔ گرادیان (CalcGradient)، ساخت درخت (Build T)، محاسبهٔ میانگین گرادیان برگ‌ها (Calc values b_j^t)، به‌روزرسانی پیش‌بینی‌ها (Update M) و محاسبهٔ درخت مرتب (Calc ordered TS) به‌صورت زیر است:

$$\begin{aligned} &CalcGradient = O(s \times n)\\ &Build\ T = O(|C| \times n)\\ &Calc\ values\ b_j^t = O(n)\\ &Update\ M = O(s \times n)\\ &Calc\ ordered\ TS = O(N_{Ts,n} \times n) \end{aligned} \qquad (۷)$$

که تمام موارد نسبت به n که همان تعداد داده‌های آموزشی است خطی است. همچنین همان‌گونه که در [۲۴] اشاره شده، بخش اصلی محاسبات مربوط به تابع Build T می‌باشد که در آن |C| تعداد جای‌گشت‌های تصادفی ساخته‌شدهٔ اولیه برای تولید درخت‌های تصادفی براساس مجموعه‌دادهٔ آموزشی ورودی است. این مقدار برای مدل پیشنهادی با توجه به اینکه از CatBoost در خروجی مدل‌های m_1 تا m_5 استفاده‌شده (یک‌بار در مجمع نخست، یک‌بار در مجمع دوم و یک‌بار در سطح دوم) نسبت به‌تعداد ویژگی‌های اولیه مجموعه‌داده مقدار بسیار کوچکی است؛ بنابراین، پیچیدگی کل الگوریتم CatBoost خطی با ضریب کوچک و درنتیجه پیچیدگی کل مدل پیشنهادی همان ماکزیمم پیچیدگی مدل‌های m_1 تا m_5 است.

۸- نتیجه‌گیری و بحث

در این پژوهش روش یادگیری عمیق و یادگیری جمعی عمیق در رده‌بندی تصاویر برچسب‌دار پزشکی سرطان پوست به‌کار گرفته شد و عملکرد روش یادگیری جمعی عمیق مورد بررسی قرارگرفت و درنهایت یک مدل مبتنی بر یادگیری جمعی عمیق چندسطحی معرفی شد و از نظر کارایی با تعدادی از مدل‌های رایج و همچین مدل ارائه‌شده در دو پژوهش مشابه مقایسه شد.

نتایج نشان دادند که کارایی استفاده از یادگیری جمعی عمیق نسبت به استفاده از مدل‌های مستقل یادگیری عمیق بهبود می‌یابد؛ چراکه عملکرد کلی تمامی مجمع‌ها همواره برابر یا بهتر از مدل‌های مستقل بوده‌است؛ لذا استفاده از یادگیری جمعی عمیق را می‌توان به‌طور کلی توصیه کرد مگر زمانی که محدودیت جدی در منابع

[Downloaded from jsdp.rcisp.ac.ir on 2026-11-04]

[DOI: 10.61186/jsdp.21.2.79]

سخت‌افزاری وجود داشته و یا به‌احتمال زمانی‌که دقت یک مدل خاص اختلاف چشم‌گیری با دیگر مدل‌های مورداستفاده در مجمع دارد.

همچنین به‌شیوهٔ تجربی تأیید شد که به‌کارگیری یک مجمع چندسطحی به‌همراه ترکیب‌کننده‌ای هوشمند می‌تواند نتایج بهتری براساس معیار F1 به‌دست آورد، اما مشروط بر اینکه ابرپارامترهای آن به‌درستی تنظیم شوند. تعیین روشی جهت تنظیم درست این ابرپارامترها نیازمند پژوهش مجزایی است.

نتایج قابل‌اتکاتر نیازمند داده‌های بیشتر و متنوع‌تر است. درصورت دسترسی به منابع کافی، می‌توان مدل پیشنهادی این پژوهش را بر روی مجموعه‌داده‌های گسترده‌تری با رده‌های بیشتر (از جمله ISIC 2019) آموزش داد و امکان بازتولید نتایج را بررسی کرد. انتخاب بهترین ترکیب از جای‌گشت‌های ممکن انتخاب دو یا سه مدل برای قرارگیری در دو مجمع یا حتی انتخاب تعداد مجامع، می‌تواند به‌عنوان یک پژوهش جدید و در ادامه پژوهش حاضر انجام شود. همچنین به‌عنوان پیشنهاد دیگری برای پژوهش‌های آینده می‌توان به‌کارگیری الگوریتم‌های تجمیع دیگری را در هرکدام از سطوح ترکیب مشخص کرد.

۹- مراجع 9-Refrences

[1]"https://data.worldbank.org/indicator/SH.MED.PHYS.ZS." Accessed: Nov. 04, 2022. [Online]. Available: https://data.worldbank.org/indicator/SH.MED.PHYS.ZS

[2] B. Rahimi, H. Nadri, H. L. Afshar, and T. Timpka, "A systematic review of the technology acceptance model in health informatics," *Appl Clin Inform*, vol. 9, no. 03, pp. 604–634, 2018.

[3] M. E. Basiri, S. Nemati, M. Abdar, S. Asadi, and U. R. Acharrya, "A novel fusion-based deep learning model for sentiment analysis of COVID-19 tweets," *Knowl Based Syst*, vol. 228, 2021, doi: 10.1016/j.knosys.2021.107242.

[4] A. Garg and V. Mago, "Role of machine learning in medical research: A survey," *Comput Sci Rev*, vol. 40, p. 100370, 2021.

[5] A. Barragán-Montero *et al.*, "Artificial intelligence and machine learning for medical imaging: A technology review," *Physica Medica*, vol. 83, pp. 242–256, 2021.

[6] U. Kamath, J. Liu, and J. Whitaker, *Deep learning for NLP and speech recognition*, vol. 84. Cham, Switzerland: Springer, 2019.

[7] M. Tsuneki, "Deep learning models in medical image analysis," *Journal of Oral Biosciences*, vol. 64, no. 3, pp.312-320, 2022.

[8] S. Dong, P. Wang, and K. Abbas, "A survey on deep learning and its applications," *Comput Sci Rev*, vol. 40, p. 100379, 2021.

[9] R. Sun *et al.*, "Survey of Image Edge Detection," *Frontiers in Signal Processing*, vol. 2, p. 826967, 2022.

[10] F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, "Breast cancer histopathological image classification using convolutional neural networks," in *2016 international joint conference on neural networks (IJCNN)*, pp. 2560–2567, 2016.

[11] X. Dong, Z. Yu, W. Cao, Y. Shi, and Q. Ma, "A survey on ensemble learning," *Front Comput Sci*, vol. 14, no. 2, pp. 241–258, 2020.

[12] M. A. Ganaie, M. Hu, and others, "Ensemble deep learning: A review," *Engineering Applications of Artificial Intelligence*, vol. 115, p.105151, 2021.

[13] A. S. Kini *et al.*, "Ensemble deep learning and internet of things-based automated COVID-19 diagnosis framework," *Contrast Media Mol Imaging*, vol. 2022, 2022.

[14] A. K. Das, S. Ghosh, S. Thunder, R. Dutta, S. Agarwal, and A. Chakrabarti, "Automatic COVID-19 detection from X-ray images using ensemble learning with convolutional neural network," *Pattern Analysis and Applications*, vol. 24, no. 3, pp. 1111–1124, 2021.

[15] M. Dildar *et al.*, "Skin cancer detection: a review using deep learning techniques," *Int J Environ Res Public Health*, vol. 18, no. 10, p. 5479, 2021.

[16] M. A. Kassem, K. M. Hosny, and M. M. Fouad, "Skin lesions classification into eight classes for ISIC 2019 using deep convolutional neural network and transfer learning," *IEEE Access*, vol. 8, pp. 114822–114832, 2020.

[17] S. Kalouche, A. Ng, and J. Duchi, "Vision-based classification of skin cancer using deep learning," *2015, conducted on Stanfords Machine Learning course (CS 229) taught*, 2016.

[18] A. Sagar and D. Jacob, "Convolutional neural networks for classifying melanoma images," *bioRxiv*, pp. 2020–2025, 2021.

[19] "Skin Cancer: Malignant vs Benign, Kaggle," Skin Cancer: Malignant vs Benign, Kaggle,. Accessed: Jan. 01, 2022. [Online]. Available: https://kaggle.com/datasets/fanconic/skin-cancer-malignant-vsbenign

[20] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[21] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," in *European conference on computer vision*, pp. 630–645, 2016.

[22] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2818–2826, 2016.

[23] M. Al-Sarem, F. Saeed, W. Boulila, A. H. Emara, M. Al-Mohaimeed, and M. Errais,

[Downloaded from jsdp.rcisp.ac.ir on 2026-11-04]

[DOI: 10.61186/jsdp.21.2.79]

"Feature selection and classification using CatBoost method for improving the performance of predicting Parkinson's disease," *in Advances on Smart and Soft Computing*, Springer, pp. 189–199, 2021.

[24] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, "CatBoost: unbiased boosting with categorical features," *Adv Neural Inf Process Syst*, vol. 31, 2018.

[25] A. Géron, *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow*. O'Reilly Media, Inc., 2022.

[۲۶] فروتن راد جواد، حورعلی مریم، کیوان راد محمدعلی، «دادگان پرسش‌وپاسخ زبان فارسی»، پردازش علائم و داده‌ها، ۲۰ (۴)، صفحات ۱۰۷-۱۲۰، ۱۴۰۲.



پویان جابری مدرک کارشناسی خود را در رشتهٔ مهندسی کامپیوتر از مرکز آموزش عالی شهرضا و مدرک کارشناسی ارشد را در همان رشته از دانشگاه شهرکرد دریافت کرده‌است. علایق پژوهشی او پردازش سیگنال، ترکیب داده و یادگیری ماشین است.

نشانی رایانامهٔ ایشان عبارت است از:
jaberi@stu.sku.ac.ir



شهلا نعمتی مدرک کارشناسی خود را در رشتهٔ مهندسی کامپیوتر از دانشگاه شیراز در سال ۱۳۸۴، کارشناسی ارشد از دانشگاه صنعتی اصفهان در سال ۱۳۸۷ و مدرک دکتری مهندسی کامپیوتر را از دانشگاه اصفهان در سال ۱۳۹۵ دریافت کرده‌است. وی از سال ۱۳۹۶ استادیار گروه مهندسی کامپیوتر دانشگاه شهرکرد است. او چندین مقاله در زمینه‌های ترکیب داده، تحلیل احساسات و پردازش صوت‌وتصویر نوشته‌است. علایق پژوهشی او پردازش سیگنال، ترکیب داده‌ها، تحلیل احساسات و داده‌کاوی است.

نشانی رایانامهٔ ایشان عبارت است از:
s.nemati@sku.ac.ir



محمداحسان بصیری مدرک کارشناسی خود را در رشتهٔ مهندسی کامپیوتر از دانشگاه شیراز در سال ۱۳۸۵، کارشناسی ارشد را از دانشگاه اصفهان در سال ۱۳۸۷ و مدرک دکترای مهندسی کامپیوتر را از دانشگاه اصفهان در سال ۱۳۹۳ دریافت کرده‌است. وی اکنون دانشیار گروه مهندسی کامپیوتر دانشگاه شهرکرد است. ایشان مقاله‌های متعددی در زمینه‌های داده‌کاوی، تحلیل احساسات و ... نوشته‌است. علایق پژوهشی او پردازش زبان طبیعی، طراحی سامانه‌های هوشمند، تحلیل احساسات و داده‌کاوی است.

نشانی رایانامهٔ ایشان عبارت است از:
basiri@sku.ac.ir

[Downloaded from jsdp.rcisp.ac.ir on 2026-11-04]

[DOI: 10.61186/jsdp.21.2.79]