

## طبقه‌بندی تصاویر با استفاده از شبکه‌های عصبی عمیق

وحیده بابائیان<sup>۱\*</sup>، شقایق مدیری<sup>۲</sup> و سیده کوثر بهلگردی<sup>۲</sup>

۱- عضو هیات علمی گروه مهندسی کامپیوتر، دانشگاه صنعتی بیرجند، بیرجند، ایران

۲- دانش آموخته مقطع کارشناسی، گروه مهندسی کامپیوتر، دانشگاه صنعتی بیرجند، بیرجند، ایران

### خلاصه

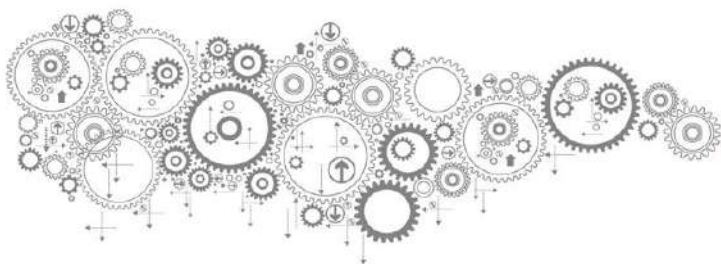
امروزه یکی از شاخه‌های وسیع و پر کاربرد هوش مصنوعی، یادگیری ماشین می‌باشد که به تنظیم و اکتشاف شیوه‌ها و الگوریتم‌ها می‌پردازد. یکی از مهمترین و جذاب ترین کاربردهای بینایی ماشین و یادگیری ماشین، طبقه‌بندی تصاویر است. از آنجایی که عمل طبقه‌بندی کاملاً وابسته به ویژگی‌های استخراج شده است، باید در استخراج و انتخاب تصاویر بسیار هوشمندانه عمل کرد تا به دقت ایده آل رسید. هدف اصلی ما در این مقاله، ارزیابی یادگیری عمیق در طبقه‌بندی تصاویر است. در این پژوهش با استفاده از کتابخانه pytorch زبان برنامه نویسی پایتون پیاده سازی انجام شده است. در طبقه بندی تصاویر، تعداد ۳۸ دسته از اشیاء در دو گروه آزمون و آموزش و مجموع ۲۷۴۶ عکس مورد بررسی قرار گرفت. به منظور ارزیابی دو روش شبکه عصبی عمیق و ماشین بردار پشتیبان مورد مقایسه قرار گرفتند. در شبکه عصبی عمیق برای آموزش از مدل از پیش آموز داده شده ResNet استفاده کرده‌ایم. نتایج ارزیابی نشان می‌دهد که طبقه‌بندی تصاویر با استفاده از روش عمیق دقت بسیار بالایی در مجموعه داده آزمون با ۸۹٪ و در مجموعه داده آموزش با ۹۵٪ را می‌دهد.

**کلمات کلیدی:** هوش مصنوعی، طبقه‌بندی تصاویر، ماشین بردار تصمیم، یادگیری عمیق، ResNet

### ۱. مقدمه

به عنوان یکی از شاخه‌های وسیع و پر کاربرد هوش مصنوعی، یادگیری ماشین به تنظیم و اکتشاف شیوه‌ها و الگوریتم‌ها می‌پردازد که بر اساس آن‌ها رایانه و سامانه‌ها توانایی تعلم و یادگیری پیدا می‌کنند. هدف یادگیری ماشین این است که رایانه بتواند به تدریج با افزایش داده‌ها کارایی بهتری در انجام وظیفه مورد نظر پیدا کند. یکی از مهمترین و جذاب ترین کاربردهای بینایی ماشین و یادگیری ماشین، طبقه‌بندی تصاویر است. از آنجایی که عمل طبقه‌بندی کاملاً وابسته به ویژگی‌های استخراج شده است، باید در استخراج و انتخاب تصاویر بسیار هوشمندانه عمل کرد تا به دقت ایده آل رسید. یادگیری عمیق باعث استخراج و انتخاب خودکار و بهینه ویژگی‌های مفید تصاویر می‌شود و همواره دقت بالا در طبقه‌بندی را به همراه دارد. برای آموزش شبکه‌های عمیق نیازمند تعداد زیادی داده ورودی هستیم. الگوریتم یادگیری همچنین می‌تواند خروجی خود را با خروجی صحیح در نظر گرفته شده مقایسه کرده و خطای خود را پیدا کند. در ادامه در بخش

\* نویسنده مسئول: [babaiyan@birjandut.ac.ir](mailto:babaiyan@birjandut.ac.ir)



دوم به مرور تحقیقات پیشین در این زمینه و در بخش سوم به پیاده سازی دو مدل در طبقه بندی تصاویر می پردازیم. بخش چهارم نتیجه گیری نهایی را بیان خواهد کرد.

## ۲. تحقیقات پیشین

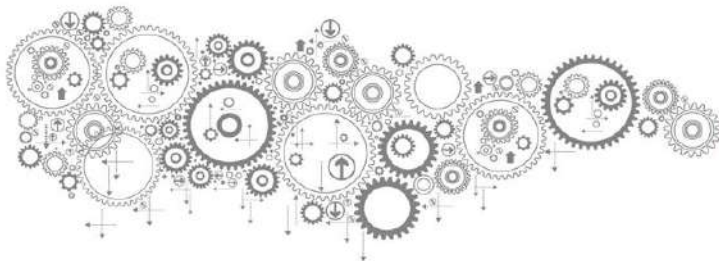
در ادبیات تحقیقات گسترده ای در زمینه طبقه بندی تصاویر در کاربردهای مختلف انجام شده است. در بین آن ها می توان به تشخیص خودکار تصاویر سونوگرافی پستان [۱]، تشخیص و طبقه بندی پوسیدگی دندان در تصاویر اشعه ایکس [۲]، طبقه بندی تومورهای مغزی [۳]، طبقه بندی تصاویر سری زمانی [۴]، طبقه بندی هسته سلول های سرطانی پستان در تصاویر هیستوپاتولوژی [۵]، تشخیص و طبقه بندی آسیب جاده ها با تصاویر تلفن های هوشمند [۶] که همگی موارد فوق با استفاده از شبکه های عصبی عمیق انجام شده است را نام برد. در سال های اخیر استفاده از روش یادگیری عمیق به عنوان یک روش مطرح با نتایج بهتر در تشخیص مورد استفاده قرار گرفته است. به عنوان مثال در ادامه نمونه هایی از تحقیقات که مبین بکارگیری و دقت بالای این الگوریتم هست را شرح می دهیم.

مقاله [۷] شامل طراحی یک سیستم دستی برای دستیابی به تصویر همراه با یک مدل یادگیری عمیق خودکار برای شناسایی سه گونه ماهی مرکب از اقیانوس آرام شمالی است که برای ارزیابی طبقه بندی سه شاخص دقت، تمرکز بر روی نقاط مشترک و میانگین زمان اجرا استفاده می شود و میانگین نتایج برای نمونه های آزمون به ترتیب ۸۵٫۷٪ و ۸۰٫۱٪ و در مدت زمان ۰٫۱۴۴ ثانیه انجام شده است. در مقاله [۸] یک رویکرد مبتنی بر یادگیری عمیق برای طبقه بندی کهکشان ها به سه دسته اصلی بیضوی، مارپیچی و نامنظم ایجاد شده است طبقه بندی تصاویر را با دقت ۹۷٫۳۹٪ طبقه بندی کرده است و بر روی ۴۶۱۴ و برای ۲۰۰ دوره آموزش داده شده است. مطالعه [۹] با استفاده از نرم افزار هوشمند برای شناسایی و طبقه بندی اشیاء و افراد در زمینه های تصویری با هدف حفاظت از حیات وحش و مدیریت را فراهم می کند. در این آزمون از ۲۰ گونه حیات وحش آفریقایی با دقت ۸۷٫۵٪ و تعداد ۱۱۱۴۶۷ استفاده شده است.

در مقاله های بررسی شده به تعداد محدودی از اشیاء که نیاز به دقت و حساسیت بیشتری دارد اشاره شده است و برای به دست آمدن نتیجه بهتر تعداد تصاویر و دفعات یادگیری به دلیل حساس بودن مسئله بیشتر است و جمع آوری تعداد زیادی تصویر از یک نوع نیازمند صرف وقت کمتری می باشد با توجه به این که هرچه تعداد تصاویر و دفعات یادگیری بیشتر باشد نتیجه بهتری حاصل می شود توصیه می شود از تعداد عکس و یادگیری بیشتری استفاده شود.

## ۳. پیاده سازی و مقایسه

در این پژوهش با استفاده از کتابخانه pytorch زبان برنامه نویسی python پیاده سازی انجام شده است [۱۰، ۱۱]. در طبقه بندی تصاویر، تعداد ۳۸ دسته از اشیاء در دو گروه آزمون و آموزش و مجموع ۲۷۴۶ عکس مورد بررسی قرار گرفت. ما از دو روش شبکه عصبی عمیق [۱۲] و ماشین بردار پشتیبان [۱۳] به منظور طبقه بندی بهره گرفته ایم. در بکارگیری شبکه عصبی عمیق برای آموزش از مدل از پیش آموز داده شده ResNet استفاده کرده ایم و به بررسی و تشخیص چند نمونه از اشیاء پرداختیم. ResNet یکی از معماری های بزرگ است که نشان می دهد یک معماری یادگیری عمیق تا چه حد می تواند عمیق باشد [۱۴] که شامل چندین ماژول رسوبی است که بر روی هم سوار شده اند و در واقع ساختمان اصلی ResNet را تشکیل می دهند. مزیت اصلی ResNet این است که هزاران هزار از این لایه های رسوبی



می توانند کنار هم قرار بگیرند تا یک شبکه را تشکیل دهند و سپس به یادگیری بپردازند. شکل ۱ چند نمونه از داده های استفاده شده در این مقاله را نشان می دهد.



شکل ۱ - نمونه داده های مورد پردازش شده

الگوریتم دیگری که در مقایسه از آن بهره می گیریم، ماشین بردار پشتیبان (SVM) می باشد چراکه در بسیاری از کاربردهای مختلف در زمینه های طبقه بندی، کارایی بالایی داشته است. SVM یک الگوریتم یادگیری ماشین است که داده ها را برای طبقه بندی، تجزیه و تحلیل می کند و یک روش یادگیری نظارت شده است که داده ها را در دسته ها مرتب می کند. ماشین بردار پشتیبان، یک دسته بند یا مرزی است که با معیار قراردادن بردارهای پشتیبان، بهترین دسته بندی و تفکیک بین داده ها را برای ما مشخص می کند. هدف این الگوریتم یافتن بهترین مرز در بین داده هاست به گونه ای که بیشترین فاصله ممکن را از تمام دسته ها داشته باشد.

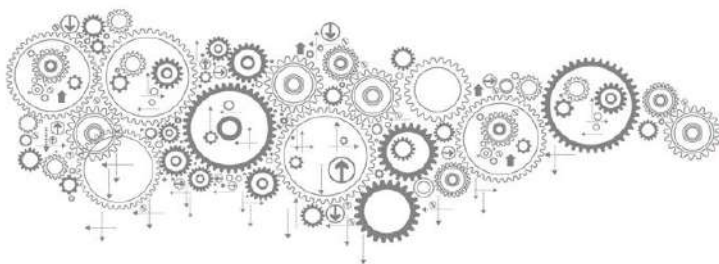
برای ارزیابی از پارامتر به عنوان معیار آماری استفاده می شود که مشخص می کند یک طبقه بندی چقدر درست تشخیص داده است. یعنی نسبت پیش بینی های صحیح (هم مثبت واقعی و هم منفی واقعی) به کل موارد بررسی شده را نشان می دهد. این پارامتر در رابطه (۱) بیان می شود:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

که در آن TP (مثبت واقعی)، FP (مثبت کاذب)، TN (منفی واقعی) و FN (منفی کاذب) می باشد. نتایج مدل سازی در جدول ۱، ارائه شده است.

جدول ۱- نتایج دقت در روش های مورد بررسی

الگوریتم	دقت	
	آزمون	آزمایش
ماشین بردار پشتیبان	٪ ۸۴	٪ ۸۸
یادگیری عمیق با شبکه عصبی کانولوشن	٪ ۸۹	٪ ۹۵



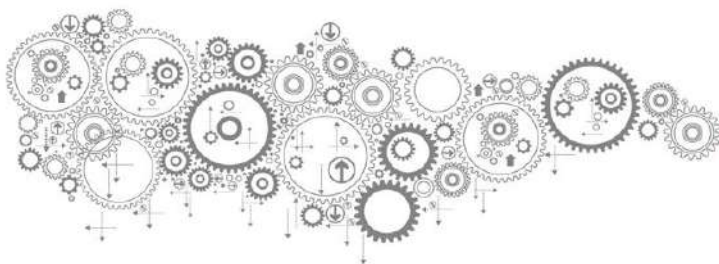
نتایج مدل سازی نشان داد یادگیری عمیق با یک شبکه عصبی کانولوشن در استخراج ویژگی بسیار مؤثرتر عمل می کند. برای مشکلات طبقه بندی تصویر، هدف این است که براساس ویژگی های عمیق سطح بالا، کدام طبقه بندی دقیق تر عمل می کند که در یادگیری عمیق با بکارگیری مدل از پیش آموزش داده شده ResNet دقت بدست آمده در مجموعه داده های آزمون ۸۹٪ و در مجموعه داده های آموزش ۹۵٪ می باشد.

#### ۴. نتیجه گیری

با گسترش روز افزون در فناوری تصویر برداری و پیشرفت های اخیر در زمینه ی اینترنت و حافظه، روزانه تصاویر بسیار زیادی تولید می شود که انواع مختلفی دارد. مدیریت دستی این تصاویر غیر ممکن است به همین دلیل برای طبقه بندی تصاویر نیازمند هوش مصنوعی و پردازش تصویر هستیم. یکی از شاخه های جذاب در هوش مصنوعی، پردازش تصویر و بینایی ماشین است. در سال های اخیر یادگیری ماشین دچار تحولات زیادی شده است. این تحول روی کار آمدن یادگیری عمیق است که ویژگی های مفید تصاویر که مورد نیاز ماست به صورت خودکار و بهینه از ورودی خام استخراج و انتخاب می شود. در کاربردهای مختلف یادگیری ماشین، بالاترین دقت ها از آن یادگیری عمیق است. هدف اصلی ما در این پروژه، ارزیابی یادگیری عمیق در طبقه بندی تصاویر است. در این پژوهش با استفاده از کتابخانه pytorch زبان برنامه نویسی پایتون پیاده سازی انجام شده است. برای آموزش از مدل از پیش آموزش داده شده ResNet استفاده کرده ایم و به بررسی و تشخیص چند نمونه از اشیا پرداختیم با تعداد ۳۸ دسته از اشیا و دو گروه آزمون و آموزش و مجموع ۲۷۴۶ داده، دقت بدست آمده در گروه اول ۸۹٪ و در گروه دوم ۹۵٪ بدست آمد. با مرور کارهای انجام شده در ادبیات و ایده گرفتن از مقاله های دیگر و همچنین پیاده سازی روش یادگیری عمیق با زبان پایتون در طبقه بندی تصاویر به این نتیجه رسیدیم که طبقه بندی تصاویر با استفاده از روش عمیق دقت بسیار بالایی را به همراه دارد. چالش های موجود در این پژوهش بالا بردن دقت طبقه بندی و جمع آوری داده های بیشتر است. می توان یک معماری جدید با عمق کمتر را با این مجموعه داده آموزش داد و نتایج آن را ارزیابی کرد و یا با استفاده از روش تقویت داده سعی در بالا بردن دقت طبقه بندی نمود.

#### ۵. مراجع

- 1.Qi, X., et al., (2019). "Automated diagnosis of breast ultrasonography images using deep neural networks". Medical image analysis, **52**: p. 185-198.
- 2.Ali, R.B., R. Ejbali, and M. Zaied. (2016). "Detection and classification of dental caries in x-ray images using deep neural networks". in *International Conference on Software Engineering Advances (ICSEA)*.
- 3.Mohsen, H., et al.,(2018). "Classification using deep learning neural networks for brain tumors". Future Computing and Informatics Journal, **3**(1): p. 68-71.
- 4.Hatami, N., Y. Gavet, and J. Debayle.(2018). "Classification of time-series images using deep convolutional neural networks". in *Tenth international conference on machine vision (ICMV 2017)*. International Society for Optics and Photonics.



5. Feng, Y., L. Zhang, and Z. Yi, (2018). "Breast cancer cell nuclei classification in histopathology images using deep neural networks". International journal of computer assisted radiology and surgery, **13**(2): p. 179-191.
6. Maeda, H., et al., (2018). "Road damage detection and classification using deep neural networks with smartphone images". Computer-Aided Civil and Infrastructure Engineering, **33**(12): p. 1127-1141.
7. Hu, J., et al., (2020). "A rapid, low-cost deep learning system to classify squid species and evaluate freshness based on digital images". Fisheries Research, **221**: p. 105376.
8. Misra, D., et al., (2020). "Convoluting cosmos: Classifying galaxy images using deep learning", in *Data Management, Analytics and Innovation*. Springer. p. 569-579.
9. Miao, Z., et al., (2019). "Insights and approaches using deep learning to classify wildlife". Scientific reports, **9**(1): p. 1-9.
10. Subramanian, V., (2018). "Deep Learning with PyTorch: A practical approach to building neural network models using PyTorch". Packt Publishing Ltd.
11. Howard, J. and S. Gugger, (2020). "Deep Learning for Coders with fastai and PyTorch". O'Reilly Media.
12. Hemanth, D.J. and V.V. Estrela, (2017). "Deep learning for image processing applications". Vol. 31. IOS Press.
13. Chandra, M.A. and S. Bedi, (2018). "Survey on SVM and their application in image classification". International Journal of Information Technology, p. 1-11.
14. Targ, S., D. Almeida, and K. Lyman, (2016). "Resnet in resnet: Generalizing residual architectures". arXiv preprint arXiv:1603.08029.