جمهوری اسلامی ایران
وزارت علوم، تحقیقات و فناوری
دانشگاه تربت حیدریه
ISC
پنجمین کنفرانس ملی
کاربرد فناوری های نوین در علوم مهندسی
5th National Conferences
Application of Novel Technologies
in Engineering Sciences
۵ و ۶ اسفندماه ۱۳۹۹
دانشگاه تربت حیدریه
www.engconfs.ir

طبقه‌بندی تصاویر با استفاده از شبکه‌های عصبی عمیق

وحیده بابائیان [*۱]، شقایق مدیری[۲] و سیده کوثر بهلگردی[۲]

۱- عضو هیات علمی‌گروه 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 می دهد که طبقه‌بندی تصاویر با استفاده از روش عمیق دقت بسیار بالایی درمجموعه 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 جذاب‌ترین 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

دوم به مرور تحقیقات پیشین در این زمینه و در بخش سوم به پیاده‌سازی دو مدل در طبقه‌بندی تصاویر می‌پردازیم. بخش چهارم نتیجه‌گیری نهایی را بیان خواهد کرد.

۲. تحقیقات پیشین

در ادبیات تحقیقات گسترده‌ای در زمینه طبقه‌بندی تصاویر در کاربردهای مختلف انجام شده است. در بین آن‌ها می‌توان به تشخیص خودکار تصاویر سونوگرافی پستان [۱]، تشخیص و طبقه‌بندی پوسیدگی دندان در تصاویر اشعه ایکس [۲]، طبقه‌بندی تومورهای مغزی[۳]، طبقه‌بندی تصاویر سری زمانی [۴]، طبقه‌بندی هسته سلول‌های سرطانی پستان در تصاویر هیستوپاتولوژی [۵]، تشخیص و طبقه‌بندی آسیب جاده‌ها با تصاویر تلفن‌های هوشمند[۶] که همگی موارد فوق با استفاده از شبکه‌های عصبی عمیق انجام شده است را نام برد. در سال‌های 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 نمونه‌هایی از تحقیقات که مبین بکارگیری و دقت بالای این الگوریتم هست را شرح می‌دهیم.

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 طبقه‌بندی سه شاخص دقت، تمرکز بر روی نقاط مشترک و میانگین زمان اجرا استفاده می‌شود و میانگین نتایج برای نمونه‌های 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 طبقه‌بندی کهکشان‌ها به سه دسته اصلی بیضوی, مارپیچی و نامنظم ایجاد شده است طبقه‌بند، تصاویر را با دقت ۹۷,۳۹٪طبقه‌بندی 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 طبقه‌بندی اشیاء و افراد در زمینه‌های تصویری با هدف حفاظت از حیات وحش و مدیریت را فراهم می‌کند. در این آزمون از ۲۰ گونه حیات وحش آفریقایی با دقت ۸۷,۵ ٪ و تعداد ۱۱۱۴۶۷ استفاده شده است.

در مقاله‌های 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 می‌باشد با توجه به این که هرچه تعداد تصاویر و دفعات یادگیری بیشتر باشد نتیجه بهتری حاصل می‌شود توصیه می‌شود 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 طبقه‌بندی بهره گرفته‌ایم.

در بکارگیری شبکه عصبی عمیق برای آموزش از مدل از پیش‌آموز داده شده ResNet استفاده کرده‌ایم و به بررسی و تشخیص چند نمونه از اشیا پرداختیم. ResNet یکی از معماری‌های بزرگ است که نشان می‌دهد یک معماری یادگیری عمیق تا چه حد می‌تواند عمیق باشد[۱۴] که شامل چندین ماژول رسوبی است که برروی هم سوار شده‌اند و در واقع ساختمان اصلی ResNet را تشکیل می‌دهند. مزیت اصلی ResNet این است که هزاران هزار از این لایه‌های رسوبی

می‌توانند کنار هم قرار بگیرند تا یک شبکه را تشکیل دهند و سپس به یادگیری بپردازند. شکل ۱ چند نمونه از داده‌های استفاده شده در این مقاله را نشان می‌دهد.

شکل ۱ - نمونه داده‌های مورد پردازش شده

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

برای ارزیابی از پارامتر به عنوان معیار آماری استفاده می‌شود 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 می‌شود:

$$Accuracy = (TP + TN) / (TP + TN + FP + FN) \quad (۱)$$

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 روش‌های 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 می‌کند. برای مشکلات طبقه‌بندی تصویر، هدف این است که براساس ویژگی‌های عمیق سطح بالا، کدام طبقه‌بندی دقیق‌تر عمل می‌کند 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٪ می‌باشد.

۴. نتیجه گیری

با گسترش روز افزون در فناوری تصویر برداری و پیشرفت‌های اخیر در زمینه ی اینترنت و حافظه، روزانه تصاویر بسیار زیادی تولید می‌شود که انواع مختلفی دارد. مدیریت دستی این تصاویر غیر ممکن است به همین دلیل برای طبقه‌بندی تصاویر نیازمند هوش مصنوعی و پردازش تصویر هستیم. یکی از شاخه‌های جذاب در هوش مصنوعی، پردازش تصویر و بینایی ماشین است. در سال‌های اخیر یادگیری ماشین دچار تحولات زیادی شده است. این تحول روی کار آمدن یادگیری عمیق است که ویژگی‌های مفید تصاویر که مورد نیاز ماست به صورت خودکار و بهینه از ورودی خام استخراج و انتخاب می‌شود. در کاربردهای مختلف یادگیری ماشین، بالاترین دقت‌ها از آن یادگیری عمیق است. هدف اصلی ما در این پروژه، ارزیابی یادگیری عمیق در طبقه‌بندی 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 آموز 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 مقاله‌های دیگر و همچنین پیاده سازی روش یادگیری عمیق با زبان پایتون در طبقه‌بندی تصاویر به این نتیجه رسیدیم که طبقه‌بندی تصاویر با استفاده از روش عمیق دقت بسیار بالایی را به همراه دارد. چالش‌های موجود در این پژوهش بالابردن دقت طبقه‌بندی و جمع‌آوری داده‌های بیشتر است. می‌توان 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 طبقه‌بندی 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). "*Convoluted 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.