Journal of Soft Computing and Information Technology (JSCIT)
Babol Noshirvani University of Technology, Babol, Iran
Journal Homepage: jscit.nit.ac.ir
Volume 13, Number 3, Autumn 2024, pp. 46-59
Received: 11/29/2024, Revised: 01/17/2025, Accepted: 01/21/2025

A Hybrid Generalized Zero-Shot Learning Method for Image Classification

Seyed Mosayeb Arabi[1], Amineh Amini[2*] and Hadi Saboohi[3]
1- Master student, Department of Computer Engineering, Karaj Branch, Islamic Azad University, Karaj, Iran.
2*- Assistant Professor, Department of Computer Engineering, Karaj Branch, Islamic Azad University, Karaj, Iran.
3- Assistant Professor, Department of Computer Engineering, Karaj Branch, Islamic Azad University, Karaj, Iran.
[1]arabi1392@gmail.com, [2*]aamini@kiau.ac.ir, and [3]saboohi@kiau.ac.ir

Corresponding author's address: Amineh Amini, Assistant Professor, Department of Computer Engineering Islamic Azad University Karaj Branch, Iran.

Abstract- One of the main challenges in machine learning is the need for extensive training data, which increases the cost and time required for data preparation. In some cases, there may be limited data available, or privacy concerns may arise. Generalized zero-shot learning (GZSL) offers a suitable option in such situations, enabling the prediction of both seen and unseen classes. Although various methods have been proposed in this field, accurately recognizing seen and unseen classes still presents multiple challenges. This paper introduces a hybrid generalized zero-shot learning (H-GZSL) approach, which improved classification accuracy in GZSL. The proposed method simultaneously utilizes two classification techniques, enhancing classification accuracy for both seen and unseen classes. To address different scenarios and resolve conflicts in classification, a correction mechanism is employed. This mechanism compares the predicted classes from the two methods and assigns the instance to the most appropriate one. The proposed method was evaluated on the AWA2, SUN, aPY, CUB, and FLOW datasets. The evaluation separately measured the performance on seen and unseen classes, and the harmonic mean of classification accuracy across these classes was calculated. Compared to previous methods, the proposed approach achieved an improvement in the harmonic mean accuracy on the CUB and SUN datasets.

Keywords- Classification, Zero-shot learning, Generalized zero-shot learning.

مجله علمی رایانش نرم و فناوری اطلاعات
دانشگاه صنعتی نوشیروانی بابل
صفحه مجله: jscit.nit.ac.ir
جلد ۱۳، شماره ۳، پائیز ۱۴۰۳، صفحه ۴۶-۵۹
دریافت: ۱۴۰۳/۰۹/۰۹، بازنگری: ۱۴۰۳/۱۰/۲۸، پذیرش: ۱۴۰۳/۱۱/۰۲

یک روش یادگیری بدون نمونه تعمیم یافته ترکیبی برای طبقه‌بندی تصاویر

سید مصیب عربی[۱]، امینه امینی*[۲]، هادی صبوحی[۳]
۱- دانشجوی کارشناسی ارشد، گروه مهندسی کامپیوتر، دانشگاه آزاد اسلامی واحد کرج، کرج، ایران.
*۲- استادیار ، گروه مهندسی کامپیوتر، دانشگاه آزاد اسلامی واحد کرج، کرج، ایران.
۳- استادیار ، گروه مهندسی کامپیوتر، دانشگاه آزاد اسلامی واحد کرج، کرج، ایران.
[1]arabi1392@gmail.com, [2*]aamini@kiau.ac.ir, [3]saboohi@kiau.ac.ir

* نشانی نویسنده مسئول: امینه امینی، کرج، بلوار موذن، دانشگاه آزاد اسلامی واحد کرج، دانشکده هوش مصنوعی.

چکیده- یکی از چالش‌های مهم در یادگیری ماشین نیاز به داده‌های آموزشی زیاد می‌باشد که باعث می‌شود تا هزینه‌ها و زمان آماده سازی این داده‌ها افزایش پیدا کند. از سویی در برخی از موارد ممکن است داده‌های زیادی در دسترس نبوده و یا مسئله حریم خصوصی وجود داشته باشد. یادگیری بدون نمونه تعمیم‌یافته توانسته گزینه مناسبی در این شرایط بوده و کلاس‌های دیده شده و دیده نشده را پیش بینی نماید. با آنکه روش‌های مختلفی در این زمینه ارائه شده است اما دقت تشخیص کلاس‌های دیده شده و دیده نشده با چالش‌های مختلفی روبه‌رو می‌باشد. در این مقاله یادگیری بدون نمونه تعمیم یافته ترکیبی (H-GZSL) معرفی شده که توانسته باعث افزایش دقت طبقه‌بندی در یادگیری بدون نمونه تعمیم‌یافته شود. در روش پیشنهادی از دو روش طبقه‌بندی به طور هم‌زمان استفاده می‌شود که باعث شده تا دقت طبقه‌بندی نمونه‌ها در کلاس‌های دیده شده و دیده نشده افزایش پیدا کند. برای مواجهه با سناریوهای مختلف و تناقض در طبقه‌بندی از یک ساز و کار اصلاح استفاده شده است. این ساز و کار کلاس پیش‌بینی شده دو روش را مقایسه و نمونه را در مناسب‌ترین کلاس طبقه‌بندی می‌نماید. روش پیشنهادی بر روی مجموعه داده‌های AWA2، SUN، aPY، CUB و FLOW مورد ارزیابی قرار گرفت. ارزیابی کلاس‌های دیده شده و نشده به طور جدا محاسبه و در نهایت میانگین هارمونیک طبقه‌بندی نمونه‌ها محاسبه گردید. روش پیشنهادی توانسته تا در مقایسه با روش‌های پیشین دقت میانگین هارمونیک را در مجموعه داده CUB و SUN افزایش دهد.

واژه‌های کلیدی: *طبقه‌بندی، یادگیری بدون نمونه، یادگیری بدون نمونه تعمیم یافته*

۱- مقدمه

یادگیری بدون نمونه (ZSL[۱]) و یادگیری بدون نمونه تعمیم‌یافته (GZSL[۲]) از جمله روش‌هایی می‌باشد که کار طبقه بندی را انجام می‌دهند. در یادگیری بدون نمونه مجموعه داده به دو کلاس دیده شده و دیده نشده تقسیم می‌شود اما تنها از کلاس‌های دیده شده برای آموزش استفاده می‌شود و از کلاس‌های دیده نشده برای آزمایش استفاده می‌شود. در یادگیری بدون نمونه تعمیم‌یافته از فضای جستجوی شامل تمام کلاس‌های دیده شده و دیده نشده برای آزمایش استفاده می‌شود[۱]. در یادگیری بدون نمونه مدل آموزش می‌بینند تا بتواند طبقه بندی یا پیش بینی در مورد کلاس‌ها یا دسته‌هایی را انجام دهد که در مرحله آموزش دیده نشده‌اند. برای این منظور از اطلاعات فرامتنی یا ویژگی‌های توصیفی مانند ویژگی‌های معنایی استفاده می‌شود که رابطه‌ای بین کلاس‌های دیده شده در داده‌های آموزشی و کلاس‌های دیده نشده برقرار می‌کند در یادگیری بدون نمونه تعمیم‌یافته مدل نه تنها باید قادر باشد کلاس‌های دیده نشده را شناسایی کند، بلکه باید بتواند بین کلاس‌های دیده شده و دیده نشده تفاوت قائل شود. این به معنای آن است که مدل باید بتواند در یک محیط واقعی که هم کلاس‌های

دیده شده و هم کلاس‌های دیده نشده وجود دارند، عملکرد خوبی داشته باشد. در حالی که یادگیری بدون نمونه استاندارد تنها بر روی کلاس‌های دیده نشده تمرکز دارد[۲]. به طور کلی یادگیری بدون نمونه تعمیم‌یافته پیچیده‌تر است، زیرا نیاز دارد که مدل بتواند در مقابل تمامی کلاس‌های دیده شده و دیده نشده عملکرد مناسبی داشته باشد[۳]. در یادگیری بدون نمونه تعمیم بافته فعالیت‌هایی که باید شناسایی شوند شامل نمونه‌های تحت پوشش داده‌های آموزشی و همچنین نمونه‌های دیده نشده می‌باشد[۴]. روش‌های مختلفی برای یادگیری بدون نمونه تعمیم‌یافته وجود دارد که هر کدام چالش‌ها و دقت طبقه بندی مخصوص به خود را دارا می‌باشد. یادگیری بدون نمونه و یادگیری بدون نمونه تعمیم‌یافته با چالش عدم تعادل داده‌ها در کلاس‌های دیده شده و نادیده مواجه هستند و مدل آموخته شده تمایل دارد تا داده‌های متعلق به کلاس‌های دیده شده را بیش از حد برازش نماید و باعث می‌شود تا مدل ارائه شده برای کلاس‌های دیده نشده عملکرد مناسبی نداشته باشد [۵].

در طبقه بندی نمونه‌ها چالش‌های زیادی وجود دارد از جمله چالش هابنس[۳] که تعداد دفعاتی می‌باشد که یک نمونه در کلاس‌های همسایه ظاهر می‌شود. چالش دیگری که وجود دارد مربوط به دامنه نمونه‌ها می‌باشد، به طور مثال دو نمونه از کلاس دیده شده شامل شیر و دلفین باشد و نمونه بعدی یک پرنده باشد. بنابراین لازم است تا با تغییر در روش‌های ارائه شده و پیشنهاد روش‌های جدید دقت یادگیری بدون نمونه تعمیم‌یافته افزایش پیدا کند. در یادگیری بدون نمونه تعمیم یافته، کاربردهای بسیاری از جمله تحلیل احساسات، طبقه بندی تصویر، تشخیص ناهنجاری‌ها، کشاورزی هوشمند وجود دارد. تشخیص اشیا و تصاویر تا کنون بر اساس بانک جامع داده‌های برچسب گذاری شده انجام شده است. برچسب گذاری این داده‌ها خود نیاز به زمان و هزینه بسیار زیادی می‌باشد که سبب شده تا طبقه بندی داده‌ها بدون نمونه آموزشی مورد توجه قرار بگیرد. تمرکز ما در این پژوهش بر روی طبقه بندی تصویر و رفع چالش‌های موجود و افزایش دقت در این زمینه می‌باشد [۶]. در سال‌های اخیر روش‌های بسیاری در مورد طبقه بندی بدون نمونه پیشنهاد شده است. در طبقه بندی بدون نمونه با آنکه آموزش مدل روی مجموعه‌ای از نمونه‌ها انجام می‌شود اما سپس می‌تواند نمونه‌های جدید از کلاس‌های دیده نشده قبلی را طبقه بندی کند. با توسعه یادگیری بدون نمونه تشخیص تصاویر آموزشی از هر دو کلاس دیده شده و دیده نشده چالش بسیار مهمی است. در طبقه بندی بدون نمونه تعمیم‌یافته مدل می‌تواند نمونه‌های جدید را هم از کلاس‌های دیده نشده و هم از کلاس‌های دیده شده طبقه بندی نماید. در پژوهش‌های اخیر با مطالعه روش‌های پیشین و بررسی چالش‌هایی که در آنها وجود داشته است سعی شده تا با ارائه راهکاری دقت را افزایش دهند.

یکی از چالش‌های مهم یادگیری بدون نمونه تعمیم‌یافته سوگیری و تعصب بسیار زیاد به کلاس‌های دیده شده در آموزش می‌باشد که روش‌های معرفی شده با تمرکز بر روی آن، دقت تشخیص یادگیری بدون نمونه تعمیم‌یافته را نسبت به روش‌های قبلی بسیار افزایش داده‌اند. در بحث یادگیری بدون نمونه تعمیم‌یافته دقت طبقه بندی در کلاس‌های دیده شده و دیده نشده جداگانه محاسبه و در نهایت با میانگین هارمونیک دقت روش پیشنهادی مشخص می‌شود [۷]. با اینکه میزان دقت در طبقه بندی تصاویر با یادگیری بدون نمونه تعمیم‌یافته طی سال‌های گذشته افزایش یافته و روش‌های مختلفی در این زمینه ارائه شده، اما دقت طبقه بندی بالایی ندارند. با توجه به تفاوت‌هایی که روش‌های پیشنهادی در این زمینه دارا می‌باشند باعث شده تا علاوه بر تفاوت دقت در مجموعه داده‌های مختلف، در طبقه بندی کلاس‌های دیده شده و دیده نشده نیز دارای دقت‌های مختلفی باشند.

برخی از روش‌ها توانسته‌اند در طبقه بندی کلاس‌های دیده شده دقت بالاتری داشته باشند و برخی دیگر در طبقه بندی کلاس‌های دیده نشده دقت بیشتری دارند. برخی از روش‌ها مبتنی بر فضای ویژگی مشترک بوده و هدف این روش‌ها یادگیری فضایی مشترک است که در آن کلاس‌های دیده شده و دیده نشده بتوانند به صورت معنایی نمایش داده شوند. در این روش‌ها ویژگی‌های کلاس به گونه‌ای نمایش داده می‌شود تا کلاس‌های دیده نشده قابل شناسایی باشد. بخش دیگری از روش‌ها مبتنی بر یادگیری متقابل بوده و اطلاعاتی چون ویژگی‌های بصری و تعبیه‌های معنایی به صورت مشترک یاد گرفته می‌شوند تا در نهایت مدل بتواند کلاس‌های دیده نشده را با ترکیب این دو اطلاعات شناسایی نماید. دسته‌ای دیگر از روش‌ها مبتنی بر انتقال دانش می‌باشد و بر روی انتقال دانش از کلاس‌های دیده شده به کلاس‌های دیده نشده تمرکز دارند. این کار اغلب با یادگیری ویژگی‌هایی انجام می‌شود که بین کلاس‌های مختلف مشترک است. همچنین برخی از روش‌ها مبتنی بر یادگیری چندکاره بوده و از آن برای تقویت تعمیم پذیری استفاده می‌کنند. به همین دلیل مدل‌ها به صورت همزمان چندین وظیفه را آموزش دیده و با کمک آن ویژگی‌های مشترک و معنایی بیشتری را می‌آموزند. بخش دیگری از روش‌ها نیز مبنی بر یادگیری تقویتی بوده و مدل‌ها برای ارتباط بهتر بین کلاس‌های دیده شده و دیده نشده از یادگیری تقویتی استفاده می‌کنند[۸]. برخی از روش‌های ارائه شده در تشخیص کلاس‌های دیده شده و برخی دیگر در تشخیص کلاس‌های دیده نشده دقت بیشتری دارند. به همین دلیل

ارائه روشی که بتواند به طور همزمان در تشخیص کلاس‌های دیده شده و دیده نشده دقت مناسبی داشته باشد اهمیت پیدا می‌کند.

یکی از چالش‌های اصلی در یادگیری ماشین نیاز به حجم زیادی از داده‌های آموزشی می‌باشد که باعث شده تا بسیاری از مدل‌های یادگیری عمیق برای عملکرد مناسب نیاز به داده‌های زیادی داشته باشند. با این حال در دنیای واقعی دسترسی به چنین حجم بزرگی از داده‌ها ممکن نیست. یادگیری بدون نمونه تعمیم‌یافته توانسته این مشکل را با کاهش نیاز به داده‌های آموزشی تا حدی برطرف نماید. از طرف دیگر یادگیری بدون نمونه تعمیم‌یافته کاربردهای گسترده‌ای در حوزه‌های مختلف دارد. برای مثال در تشخیص پزشکی که داده‌های آموزشی کمی برای یک بیماری نادر وجود دارد اهمیت یادگیری بدون نمونه تعمیم‌یافته مشخص می‌شود[۹]. یکی از ویژگی‌های مهم یادگیری بدون نمونه تعمیم‌یافته، توانایی تطبیق با شرایط جدید بوده که باعث می‌شود بتواند در حوزه‌هایی که داده‌های جدید به سرعت تولید می‌شوند گزینه مناسبی باشد. یادگیری بدون نمونه تعمیم‌یافته به مدل‌ها کمک می‌کند تا با استفاده از داده‌های کمتر به یادگیری بهتری دست پیدا کنند. از این رو توانایی یادگیری مستقل مدل‌ها افزایش یافته و با کمک تجربیات قبلی و یادگیری انتقالی قادر به حل مسائل جدید می‌باشند. این ویژگی در توسعه مدل‌های هوش مصنوعی که لازم است بتوانند در شرایط واقعی و با داده‌های محدود به خوبی عمل کنند بسیار اهمیت دارد.

در این مقاله، با ارائه روش یادگیری بدون نمونه تعمیم‌یافته ترکیبی با تکیه بر طبقه بندی همزمان توسط دو روش مختلف، دقت تشخیص افزایش پیدا کرده است تا همزمان چالش کلاس‌های دیده شده و دیده نشده به حداقل برسد. در روش پیشنهادی زمانی که کلاس طبقه بندی هر دو روش یکسان باشد، نمونه در کلاس مربوطه قرار می‌گیرد. همچنین برای مواقعی که کلاس پیش بینی شده توسط دو روش متفاوت باشد ساز و کار اصلاح پیشنهاد شده است. ساز و کار اصلاح، برای حل تناقض کلاس پیش بینی شده میزان شباهت نمونه را با یک نمونه از هر کدام از کلاس‌های پیش بینی شده انجام می‌دهد. نمونه مد نظر در کلاسی که بیشترین شباهت را دارد، طبقه بندی می‌شود. در صورت عدم وجود حداقل یک نمونه در کلاس پیش بینی شده، ساز و کار اصلاح اطلاعات لازم مانند نمونه و دو کلاس پیش بینی شده را ذخیره می‌نماید. با طبقه بندی هر نمونه، ساز و کار اصلاح شرایط نمونه‌های مد نظر را برای مقایسه میزان شباهت بررسی می‌نماید و در صورت وجود شرایط نمونه را بر اساس میزان شباهت با نمونه‌های هر دو کلاس پیش بینی شده، طبقه بندی می‌نماید.

ساختار مقاله به صورت زیر است: در بخش ۲ روش‌های ارائه شده یادگیری بدون نمونه تعمیم‌یافته مرور می‌شود. در بخش ۳ چالش‌های یادگیری بدون نمونه تعمیم‌یافته بررسی می‌شود. در بخش ۴ روش پیشنهادی یادگیری بدون نمونه تعمیم‌یافته ترکیبی معرفی می‌شود. در بخش ۵ نتایج روش پیشنهادی ارزیابی و مورد بحث قرار می‌گیرد. در بخش ۶ نتیجه گیری شده و در بخش ۷ نیز پیشنهاداتی برای کارهای آینده ارائه می‌شود.

۲- کارهای مرتبط

در یادگیری بدون نمونه تعمیم‌یافته همزمان باید کلاس‌های دیده شده و دیده نشده تشخیص داده شوند و روش‌های مختلفی در این زمینه معرفی شده است. در مقاله [۱۰] که در سال ۲۰۱۳ ارائه شد یک مدل یادگیری ماشین معرفی شده است که قادر است اشیاء را در تصاویر شناسایی کند حتی اگر داده‌های آموزشی برای کلاس آن شیء وجود نداشته باشد. مقاله [۱۱] رویکرد استفاده از پشته‌های استاندارد شده را پیشنهاد می‌دهد. این رویکرد با تنظیم نتایج تابع‌های تشخیصی برای کلاس‌های دیده شده و دیده نشده، دقت شناسایی را بهبود می‌بخشد. در مقاله [۱۲] رویکرد تولید داده‌های آموزشی مصنوعی برای دسته‌های بدون نمونه پیشنهاد شده است. در مقاله [۱۳] با استفاده از طبقه‌بندی سلسله مراتبی، برای هر شیء از یک کلاس ناشناخته نزدیک‌ترین کلاس فوقانی در سلسله مراتب کلاس‌های شناخته شده پیدا می‌شود.

مقاله [۱۴] به بررسی مشکل عدم تعادل شدید در داده‌های آموزشی بین کلاس‌های دیده شده و دیده نشده پرداخته و برای رفع نیاز به نمونه‌های برچسب‌دار از کلاس‌های دیده نشده، یک شبکه مولد تخاصمی (GAN[۴]) جدید را پیشنهاد می‌دهد که ویژگی‌های CNN را بر اساس اطلاعات معنایی در سطح کلاس تولید می‌کند. روش پیشنهادی مقاله [۵] ترکیبی از یک GAN است که توانایی تولید ویژگی‌های CNN تمایز پذیر را برای آموزش طبقه‌بندی‌های softmax یا هر روش جاسازی چندوجهی دارد. هاوفنگ ژانگ در مقاله [۱۵] روش جدیدی برای چالش فرا تناسب کلاس (CO[۵]) و تناسب بیش از حد مدل با کلاس‌های دیده شده و ناتوانی در تعمیم به کلاس‌های دیده، معرفی کرده است.

در مقاله [۱۶] روشی برای حفظ ساختار فضای معنایی با استفاده از روابط معنایی بین دسته‌ها پیشنهاد شده است که معنایی بودن را به فضای تعبیه القا می‌کند. بر خلاف کارهای قبلی، در مقاله [۱۷] تعبیه‌ای با بعد پایین ایجاد شده که به صورت بصری معنایی است. رویکرد پیشنهادی از یک مدل گرافیکی سه گره‌ای بهره می‌برد که

شامل متغیرهای معنایی (S)، ورودی (X) و برچسب (Y) است. همچنین مقاله [۱۸] به معرفی یک مدل مولد مبتنی بر یادگیری متا برای یادگیری بدون نمونه می‌پردازد. این رویکرد از یادگیری متا برای آموزش یک مدل مولد عمیق استفاده می‌کند که رمز گذار خودکار متغیر (VAE[۶]) و شبکه‌های مولد تقابلی (GAN) را ترکیب می‌کند. مقاله [۱۹] یک شبکه انتشار ویژگی‌ها (APNet[۷]) پیشنهاد داده است که با بهره‌گیری از روابط بین کلاسی، بردار ویژگی‌های هر کلاس را بهبود بخشیده و به دسته‌بندی نزدیک‌ترین همسایه (KNN[۸]) برای پیش‌بینی تصاویر با استفاده از بردار ویژگی‌های اصلاح شده تکیه می‌کند.

جی شاو و همکاران در مقاله [۲۰] مدل جدیدی به نام MCGM-VAE[۹] را با هدف یافتن توزیع احتمالی بهینه برای نمایش‌های نهان پیشنهاد کرده است تا فاصله بین ویژگی‌های بصری و معنایی پس از کدگذاری به حداقل برسد. مقاله [۲۱] روشی برای یادگیری ویژگی‌های نامرئی پیشنهاد شده است که به کلاس‌های دیده نشده کمک می‌کند تا به درستی طبقه‌بندی شوند. در مقاله [۲۲] مدل پیشنهادی بر اساس ترکیب شبکه‌های مولد متضاد (GAN) و خود رمز گذار متغیر (VAE) طراحی شده است که در آن از اطلاعات تمایز دهنده ویژگی‌های بصری استفاده می‌شود. مقاله [۲۳] برای کاهش تداخل توزیع بین ویژگی‌های تولید شده از دسته‌های دیده شده و دیده نشده، روشی شامل دو خود رمز گذار متغیر (VAE) با یک تابع ضرر سه‌گانه (triplet loss) برای یادگیری ویژگی‌های نهان تمایز بین‌کلاسی ارائه کرده است.

در مقاله [۲۴] روش SELAR[۱۰] معرفی شده است که یک روش ساده و مؤثر برای تقویت مکانیابی ویژگی‌ها است. SELAR با استفاده از روش تجمیع ماکزیمم (GMP[۱۱]) به جای تجمیع میانگین (GAP[۱۲]) در شبکه‌های عصبی کانولوشن معمولی، مکان یابی ویژگی‌ها را تقویت می‌کند. مقاله [۲۵] با استفاده از یک ساز و کار کدگذاری معنایی و یک تشخیص خارج از توزیع (OOD[۱۳])، از داده‌های معنایی و بصری برای تشخیص دقیق دسته‌های جدید استفاده می‌کند. در مقاله [۲۶] روشی با استفاده از یک استراتژی انتقال و افزایش برای تولید نمونه‌های مجازی و بهبود فرآیندهای یادگیری توسعه داده شده است. مرحله انتقال از دو منبع شامل دانش رابطه‌ای محلی (LRK[۱۴]) و دانش رابطه‌ای جهانی (GRK[۱۵]) تشکیل شده است. مقاله [۲۷] یک روش جدید با استفاده از شبکه عصبی جدید مبتنی بر خود رمز گذار متغیر واسرشتاین (WAE[۱۶]) ارائه کرده است.

در مقاله [۷] مدل پیشنهادی به نام یادگیری تفکیکی شبکه‌های مولد تقابلی طراحی شده است که ویژگی‌های معنایی واقعی و پنهان را در یک چارچوب GAN ادغام می‌کند تا از طبقه‌بندی‌کننده‌های مبتنی بر اطمینان برای تشخیص کلاس‌های دیده شده و دیده نشده استفاده کند. یک مدل جدید برای یادگیری بدون نمونه تعمیم‌یافته به نام BMCoGAN در مقاله [۲۸] معرفی شده که از نگاشت دوطرفه و شبکه‌های مولد تقابلی (GAN) استفاده می‌کند تا توزیع مشترک کلاس‌های دیده و دیده نشده را یاد بگیرد و ویژگی‌های تبعیض‌آمیز بین این کلاس‌ها را حفظ کند. مقاله [۲۹] یک روش به نام شبکه‌های مولد تقابلی چندگرافی تطبیقی (MGA-GAN) برای حل مسائل مرتبط با یادگیری بدون نمونه پیشنهاد می‌دهد. مقاله [۸] یک معماری جدید به نام جداسازی محتوا-ویژگی را به منظور بهبود قابلیت تعمیم و دقت در شناسایی کلاس‌های دیده نشده طراحی کرده است.

یک چارچوب جدید به نام کاهش ابهام ویژگی‌های دو طرفه همراستا (DFCA[۱۷]) در مقاله [۳۰] با هدف بهبود دقت در شناسایی دسته‌های دیده نشده و دیده شده و کاهش ابهام در پیش‌بینی این دسته‌ها طراحی شده است. روش DFCA با استفاده از شبکه‌های عصبی برگشت‌پذیر شرطی (cINN[۱۸]) به تولید ویژگی‌های بصری معتبر و متمایز برای دسته‌های ندیده می‌پردازد . زیانگ وانگ نیز در مقاله [۳۱] یک چارچوب نوین به نام تعبیه پیکسلی تقویت شده با زبان (LAPE[۱۹]) برای یادگیری بدون نمونه تعمیم‌یافته ارائه می‌دهد. مقاله [۳۲] روشی برای تولید ویژگی‌های تعصبی-آگاه برای یادگیری بدون نمونه تعمیم‌یافته معرفی می‌کند که به نام ABA-GAN شناخته می‌شود. همچنین روشی به نام تقویت بخش‌های دامنه‌ناپذیر (EMP[۲۰]) برای بهبود دقت در شناسایی کلاس‌های دیده نشده در مقاله [۳۳] معرفی شده است. مقاله [۳۴] روشی به نام یادگیری نمایش‌های نهفته‌ی همسان ساز مدالیته یا LCR-GAN را معرفی کرده است. این روش با تعبیه‌ی ویژگی‌های بصری و ویژگی‌های معنایی در یک فضای نهفته‌ی مشترک و با استفاده از محدودیت‌های مبتنی بر اطلاعات متقابل به دنبال حل مشکل ناهماهنگی بین ویژگی‌های بصری و معنایی هستند.

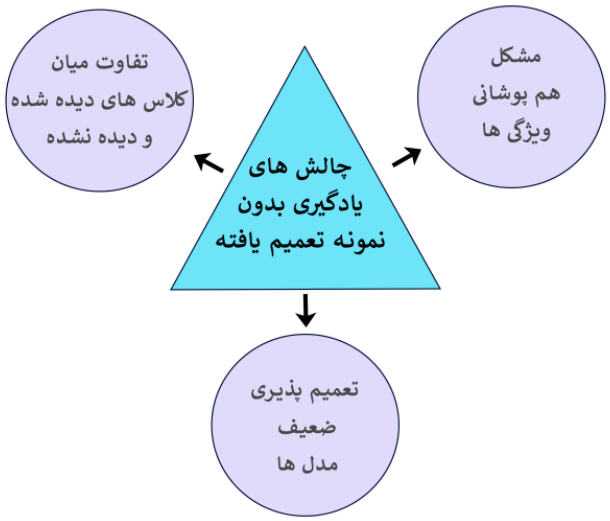
مقاله [۳۵] رویکردی جدید برای یادگیری بدون نمونه تعمیم‌یافته ارائه داده تا چالش اصلی شناسایی دسته بندی‌هایی که داده‌های بصری برای آنها در دسترس نیست را برطرف نماید. یک روش به نام شبکه‌ی مولد خود نظارتی تقویت شده با پروتوتایپ (PSGN[۲۱]) در زمینه یادگیری بدون نمونه تعمیم‌یافته در مقاله [۳۶] ارائه شده است. روش پیشنهادی PSGN با تلفیق یادگیری خودنظارتی و یادگیری پروتوتایپ، مدلی برای تولید ویژگی‌ها در GZSL ارائه می‌کند. در مقاله [۳۷] شبکه نوآورانه‌ای به نام شبکه انزوایی و تقطیر

(IDN[۳۲]) برای یادگیری بدون نمونه تعمیم‌یافته معرفی شده است. هدف این مدل بهبود عملکرد شناسایی کلاس‌های دیده نشده با استفاده از نمونه‌های دیده شده و اطلاعات توصیفی است.

۳- چالش‌های یادگیری بدون نمونه تعمیم‌یافته

در سال‌های اخیر یادگیری بدون نمونه تعمیم‌یافته با چالش‌های متعددی مواجه شده است که مانع از دست یابی به دقت بالا و کارایی مطلوب در بسیاری از کاربردهای عملی می‌باشد. این چالش‌ها در شکل ۱ نشان داده شده و شامل موارد زیر هستند:

۱. تفاوت میان کلاس‌های دیده شده و دیده نشده: یکی از مهم ترین چالش‌ها در GZSL تفاوت‌های بین کلاس‌های دیده شده مدل در طی آموزش و کلاس‌های دیده نشده می‌باشد. این تفاوت‌ها می‌توانند باعث بروز سوگیری و تعصب در مدل به سمت کلاس‌های دیده شده شوند که در نتیجه منجر به کاهش دقت در تشخیص کلاس‌های دیده نشده می‌گردد.



شکل ۱: چالش‌های یادگیری بدون نمونه تعمیم‌یافته

۲. مشکل هم پوشانی ویژگی‌ها: در بسیاری از موارد ویژگی‌های کلاس‌های مختلف ممکن است هم پوشانی داشته باشند یا مشابه باشند که این امر تشخیص دقیق کلاس‌های دیده نشده را دشوار می‌کند. مدل‌ها ممکن است نتوانند به درستی تمایز و تفاوت بین ویژگی‌های مربوط به کلاس‌های مختلف را بیاموزند [۳۸]

۳. تعمیم پذیری ضعیف مدل‌ها: مدل‌های GZSL غالباً با مشکل تعمیم پذیری مواجه هستند. به همین دلیل ممکن است مدل‌ها در شرایط آزمایشی به خوبی عمل کنند اما در مواجهه با داده‌های واقعی و پیچیده تر عملکرد آنها کاهش یابد.

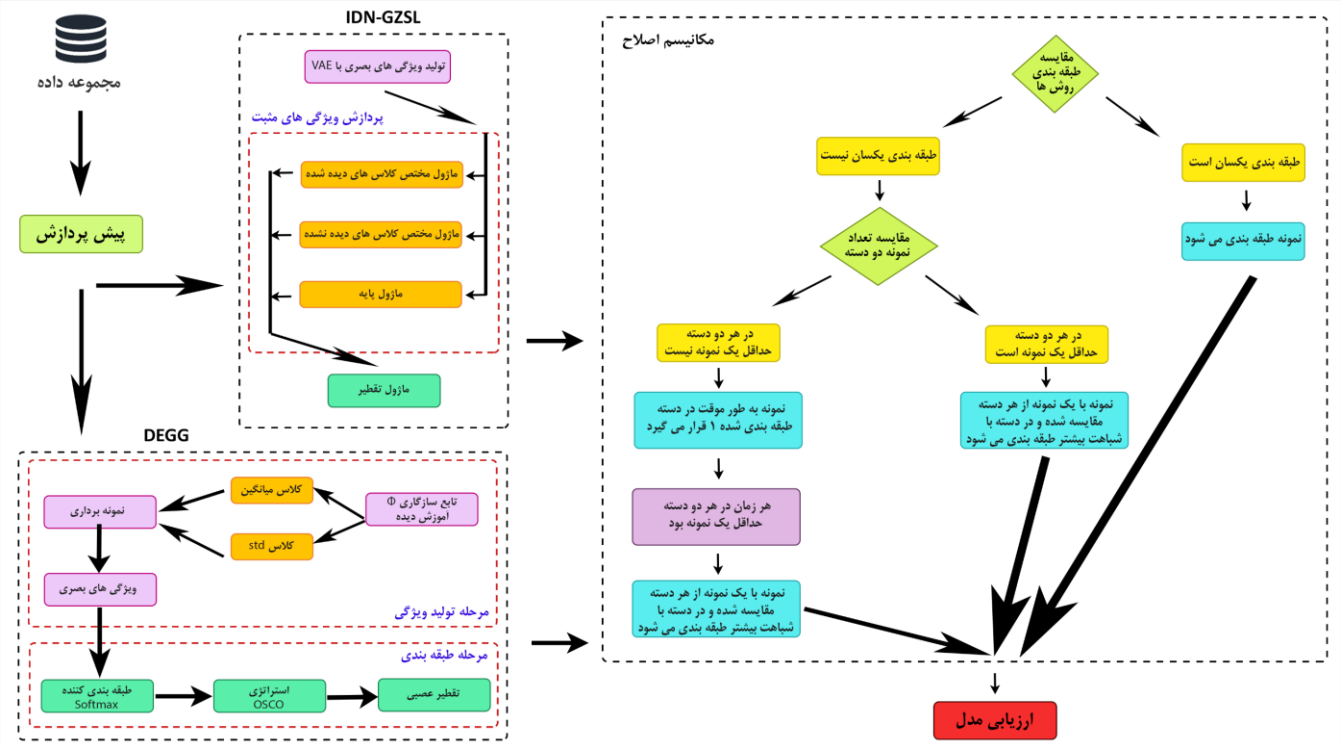
برای مقابله با این چالش‌ها سعی شده در روش‌های جدید راهکارهای مختلفی مورد استفاده قرار گرفته است. استفاده از شبکه‌های عصبی مبتنی بر توجه باعث شده تا همبستگی بین ویژگی‌های بصری و معنایی افزایش پیدا نماید. از این رو ویژگی‌های کلاس‌های دیده شده و دیده نشده به طور موثر تری شناسایی و تفکیک می‌شوند. ضمنا استفاده از روش‌های کالیبراسیون چون شبکه‌های کالیبراسیون عمیق با تنظیم بهتر پارامترهای مدل می‌تواند تعادل بهتری را بین کلاس‌های دیده شده و دیده نشده ایجاد نماید. همچنین با راه حل تولید نمونه‌های مصنوعی از کلاس‌های دیده نشده و ترکیب آن با داده‌های واقعی مدل می‌تواند در مواجه با داده‌های جدید عملکرد بهتری داشته باشد. شناخت این چالش‌ها کمک می‌کند تا برای افزایش دقت طبقه بندی تصاویر بتوان بهتر عمل نمود [۳۸].

۴- روش یادگیری بدون نمونه تعمیم‌یافته ترکیبی H-GZSL

با بررسی روش‌های موجود در زمینه یادگیری بدون نمونه تعمیم‌یافته با تمرکز بر قسمت‌های مختلف سعی بر مواجه با چالش‌های مطرح شده و در نهایت افزایش دقت طبقه بندی دارند. یکی از مهم ترین نقاط ضعف در این زمینه تمرکز روش‌های ارائه شده در یک حوزه بوده که باعث می‌شود تا همه چالش‌ها به طور هم زمان بررسی و برطرف نشوند. بنابراین روش پیشنهادی یادگیری بدون نمونه تعمیم‌یافته هیبردی که در پژوهش با نام H-GZSL برده می‌شود، با تکیه بر دو روش مختلف انجام می‌شود. این کار باعث می‌شود تا بر چالش‌های بیشتری توجه داشته و در نهایت باعث افزایش دقت طبقه بندی تصاویر شود. دو مقاله [۳۷] و مقاله [۳۵] به کار ما مرتبط بوده و با توجه به تمرکز هر کدام بر روی کلاس‌های دیده شده و دیده نشده، می‌توانند به افزایش دقت طبقه بندی کمک نمایند

روش ارائه شده در مقاله [۳۷] به طور کامل بر اساس ایجاد شبکه‌ای متشکل از سه ماژول جداگانه طراحی شده است که هدف آن‌ها بهبود دقت طبقه بندی تصاویر در زمینه یادگیری بدون نمونه تعمیم‌یافته است. سه ماژول مجزا برای پردازش ویژگی‌های مثبت در سه سناریوی مختلف طراحی شده است:

- ماژول مختص کلاس‌های دیده شده: این ماژول بر روی کلاس‌های دیده شده تمرکز دارد و به تقویت ویژگی‌هایی می‌پردازد که در طبقه بندی کلاس‌های دیده شده کاربرد دارند.
- ماژول مختص کلاس‌های دیده نشده: این ماژول برای بهبود توانایی تعمیم دهی مدل و طبقه بندی کلاس‌های



شکل ۲: یادگیری بدون نمونه تعمیم یافته ترکیبی H-GZSL

دیده نشده طراحی شده است.

- ماژول پایه: این ماژول به عنوان بخش آموزش شبکه عمل نموده و ارتباطات لازم بین کلاس‌های دیده شده و دیده نشده را فراهم می‌کند.

این روش با استفاده از آموزش تقطیر ارتباطات غیر مستقیم بین ماژول‌ها را حفظ می‌کند تا از جدایی بیش از حد بین آن‌ها جلوگیری شود. در کنار روش معرفی شده روش دیگری نیز به نام DEGG در مقاله [۳۵] انتخاب شده که می‌تواند به عنوان مکمل روش قبلی باشد که در آن از ویژگی‌های گوسی برای بهبود دقت طبقه بندی تصاویر استفاده می‌شود. در این نوع یادگیری هدف شناسایی کلاس‌هایی است که مدل آموزشی آنها فقط بر روی کلاس‌های دیده شده آموزش دیده است. این رویکرد بر اساس تولید ویژگی‌های بصری مصنوعی از اطلاعات معنایی با استفاده از توزیع‌های گوسی به جای استفاده از شبکه‌های مولد متخاصم (GANs) پیچیده و ناپایدار است. ایده اصلی این است که از توزیع‌های گوسی برای تولید ویژگی‌های بصری استفاده شود که بتوانند ویژگی‌های واقعی را برای کلاس‌های دیده شده و دیده نشده تقلید کنند. این توزیع‌ها با استفاده از آمارهای مرتبه اول و دوم (میانگین و کوواریانس) و با کمک یک تابع سازگاری پیش آموزش داده شده تخمین زده می‌شوند. روش پیشنهادی یادگیری بدون نمونه تعمیم‌یافته ترکیبی H-GZSL، نحوه اجرای هر کدام از مدل‌ها و ساز و کار اصلاح در شکل ۲ نشان داده شده است.

برای مقابله با مسئله هابنس و جایی که کلاس‌های دیده نشده ممکن است به طور نادرست به نواحی مربوط به کلاس‌های دیده شده اختصاص داده شوند، این روش از یک مجموعه طبقه بندی کننده‌های Softmax استفاده می‌کند. هر طبقه بندی کننده با کنار گذاشتن یک کلاس دیده شده در هر نوبت آموزش داده می‌شود و کمک می‌کند تا عملکرد بین کلاس‌های دیده شده و دیده نشده متعادل شود. در نهایت برای کاهش بار محاسباتی در هنگام استنتاج روش از تقطیر عصبی استفاده می‌کند. دو روش معرفی شده می‌توانند بر بخش‌های مختلفی تمرکز داشته و هر کدام به خوبی طبقه بندی تصاویر را با دقت بسیار بالایی انجام دهند. برای ترکیب این دو روش یک ساز و کار اصلاح پیشنهاد گردید.

این ساز و کار اصلاح طبقه بندی هر کدام از روش‌ها را مقایسه نموده و در صورتی که نتیجه یکسان باشد، نمونه مورد نظر را در دسته مربوطه قرار می‌دهد. برای مواقعی که هر کدام از روش‌ها دسته‌های مختلفی را برای طبقه بندی انتخاب کنند این ساز و کار در ابتدا بررسی می‌کند تا اگر از هر دسته مورد نظر حداقل یک نمونه موجود باشد، کار مقایسه نمونه انجام می‌شود. برای این کار نمونه مد نظر با یک نمونه تصادفی از هر دو دسته مقایسه شده و در دسته‌ای که

بیشترین شباهت را دارا می‌باشد قرار می‌گیرد. این ساز و کار برای مواردی که در هر دو دسته مربوطه حداقل یک نمونه وجود نداشته باشد، نمونه را به صورت موقت در دسته‌ای که روش پیشنهادی در مقاله [۳۷] طبقه بندی کرده قرار می‌دهد. دلیل این موضوع دقت بیشتر این روش در طبقه بندی تصاویر می‌باشد. در کنار طبقه بندی موقت، نمونه مورد نظر و دو دسته پیشنهادی در فضایی ذخیره می‌شود تا هر زمان که از هر دو دسته حداقل یک نمونه وجود داشت کار مقایسه نمونه با نمونه‌های هر دسته انجام شود. شبه کد روش پیشنهادی در الگوریتم ۱ نشان داده شده است.

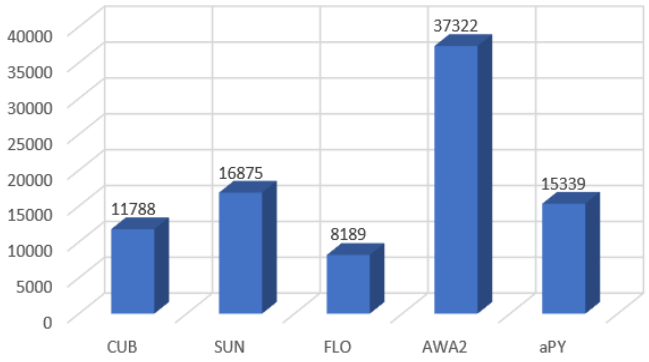
الگوریتم ۱: شبه کد یادگیری بدون نمونه تعمیم‌یافته ترکیبی H-GZSL

Algorithm1: H-GZSL Pseudocode for Hybrid Generalized Zero-Shot Learning

```
Input: Dataset
Output: Class Predictions
Perform Classification using IDN-GZSL
Perform Classification using DEGG
Compare the Results
if Classifications are Identical then
    return class label
else
    Create Different Class
    Assign to the Appropriate Category
    Check for Minimum Sample Presence
    if Minimum Sample Exists then
        Compare with Random Sample
    else
        if No Sample Exists in both Class then
            Perform Temporary Classification
            Store for Future Comparisons
        else
            Assign to the Most Similar Class
        end if
    end if
end if
```

۴-۱- مجموعه داده‌ها

روش‌های مختلفی در زمینه یادگیری بدون نمونه تعمیم‌یافته و طبقه بندی تصاویر پیشنهاد شده اند که بر روی مجموعه داده‌های مختلفی انجام شده اند. با بررسی مقالات اخیر سعی شد تا مجموعه داده به شکلی انتخاب شود که بین روش‌ها مشترک بوده و بتوان مقایسه بهتری بین روش پیشنهادی با روش‌های قبلی داشت. از این رو مجموعه داده‌های CUB، SUN، FLO، AWA2 و aPY انتخاب گردید. در شکل ۳ تعداد نمونه‌ها و تصاویر هر مجموعه داده نمایش داده شده است.



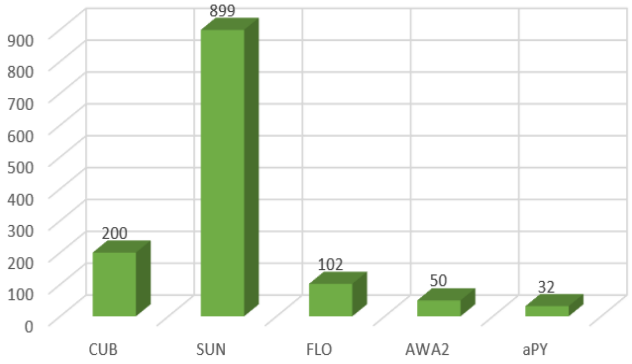
شکل ۳: تعداد تصاویر هر مجموعه داده

دیتاست CUB یکی از معروف ترین و پر استفاده ترین مجموعه داده‌ها در زمینه بینایی ماشین و یادگیری عمیق می‌باشد که شامل طیف وسیعی از تصاویر پرندگان است. تعداد تصاویر این دیتاست ۱۱۷۸۸ تصویر از پرندگان بوده که در ۲۰۰ کلاس مختلف دسته بندی شده است. هر کدام از کلاس‌ها به گونه‌ای خاص از پرندگان اختصاص دارد. هر تصویر دارای برچسب‌های متنی شامل جنسیت و سن پرندگان می‌باشد. ویژگی مهم این مجموعه داده حاشیه نویسی‌های دقیق آن می‌باشد که شامل بخش‌های مختلف بدن پرنده چون بال‌ها، منقار، دم و غیره می‌باشد. مجموعه داده دیگری که برای این پژوهش انتخاب گردید دیتاست SUN می‌باشد که مجموعه‌ای جامع از تصاویر دارای حاشیه نویسی در حوزه‌های مختلف محیطی و مکانی می‌باشد. این مجموعه داده به طور خاص برای تحقیقات در زمینه بینایی ماشین و یادگیری عمیق طراحی شده است.

این دیتاست شامل بیش از ۱۶۸۷۵ تصویر بوده که به صورت دستی حاشیه نویسی شده اند. تصاویر این مجموعه دارای ۸۹۹ دسته بندی مختلف از صحنه‌ها و محیط‌های داخلی و خارجی می‌باشد. حاشیه نویسی‌های دقیق این تصاویر باعث شده تا برای موارد طبقه بندی و تشخیص اشیاء گزینه بسیار مناسبی باشد. مجموعه داده FLO شامل تصاویر مختلفی از گل‌ها در شرایط مختلف نوری و ابعاد گل می‌باشد. این دیتاست از ۱۰۲ دسته مختلف گل تشکیل شده و تعداد تصاویر آن ۸۱۸۹ گل می‌باشد. تعداد تصاویر هر دسته تنوع زیادی در زوایا و پس زمینه دارد و همین نیز باعث چالش‌هایی در تشخیص و طبقه

بندی آن شده است. ضمنا تصاویر این مجموعه دارای حاشیه نویسی‌های بسیار دقیقی می‌باشد. یکی از مجموعه داده‌های معروف در زمینه بینایی ماشین و یادگیری عمیق دیتاست AWA2 که شامل تعداد بسیار زیادی تصویر بوده و برای شناسایی و طبقه بندی تصاویر حیوانات طراحی شده است. تعداد تصاویر این مجموعه داده شامل ۳۷۳۲۲ تصویر از حیوانات بوده و در ۵۰ کلاس مختلف از حیوانات دسته بندی شده است. هر کدام از دسته‌ها دارای تعداد زیادی تصویر بوده که دارای شرایط مختلفی می‌باشد و به شکل دقیق تری می‌تواند روش پیشنهادی را مورد ارزیابی قرار دهد. آخرین مجموعه داده‌ای که برای این پژوهش انتخاب شد دیتاست aPY می‌باشد که شامل ۱۵۳۳۹ تصاویر در سه دسته کلی حیوانات، اشیاء و وسایل نقلیه است.

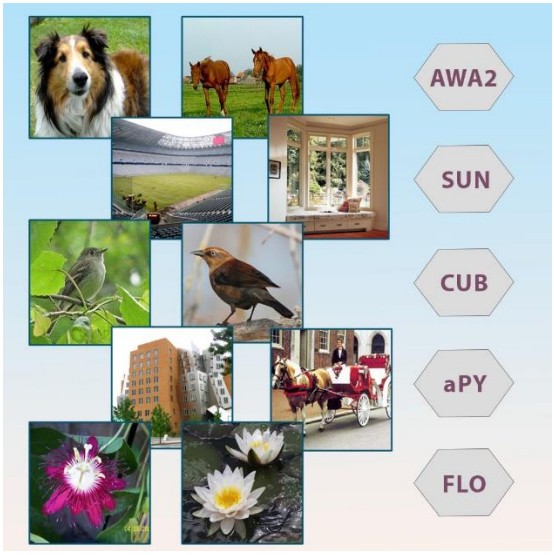
در نهایت این دیتاست از ۳۲ زیرمجموعه شامل دسته‌هایی چون هواپیما، گور خر و غیره تشکیل شده است. انتخاب این ۵ مجموعه داده باعث می‌شود تا روش پیشنهادی به خوبی ارزیابی شده و دقت آن در شرایط مختلف نمونه‌ها شامل چالش‌های زیادی چون تنوع نور، تنوع پس زمینه و نزدیکی ویژگی‌های توصیفی بررسی شود. در شکل ۴ تعداد کلاس‌های این مجموعه داده‌ها نمایش داده شده است.



شکل ۴: تعداد کلاس هر مجموعه داده

۴-۲- پیش پردازش داده‌ها

داده‌های مورد استفاده در این پژوهش به طور گسترده در تحقیقات مورد استفاده قرار گرفته و به طور معمول داده‌ها به خوبی سازماندهی و آماده سازی شده اند. در واقع داده‌ها قبل از استفاده بررسی شده اند تا مشخص شود که همه تصاویر دارای برچسب‌های صحیح و معتبر باشند. نمونه هایی از هر دیتاست در شکل ۵ قرار گرفته شده است.



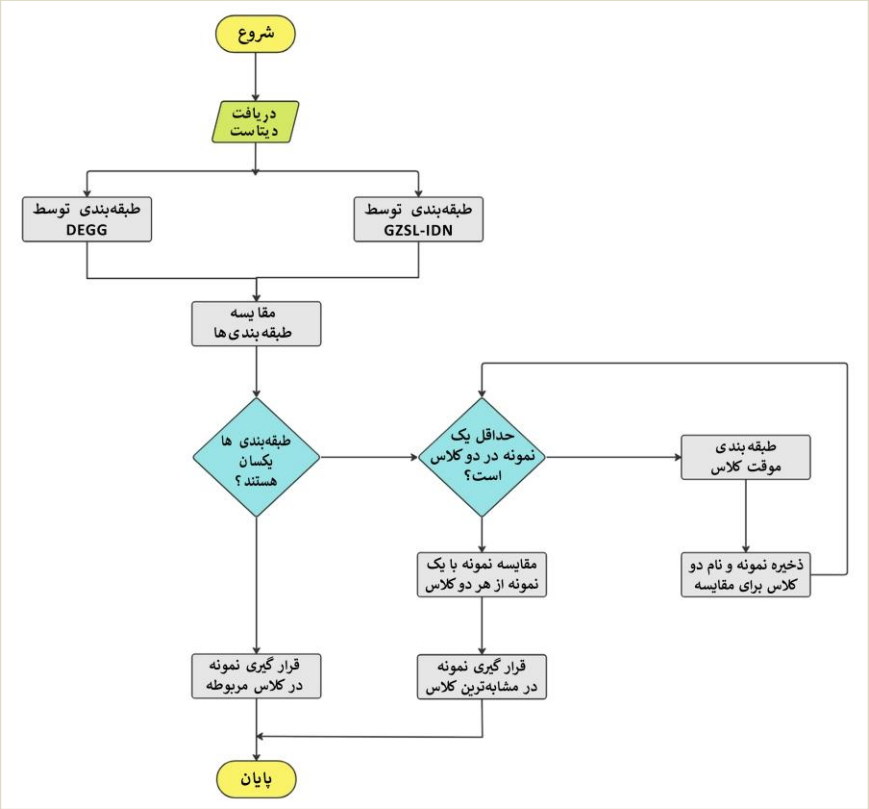
شکل ۵: نمونه هایی از دیتاست‌های مورد استفاده

۴-۳- جداسازی داده‌های آموزشی و آزمایشی

در روش‌های یادگیری بدون نمونه تعمیم‌یافته داده‌ها به مجموعه‌های آموزشی و آزمایشی تقسیم می‌شوند. در روش پیشنهادی لازم است تا پروسه آموزش و آزمایش انجام شود و به همین دلیل باید داده‌های آموزشی و آزمایشی به طور جداگانه مورد استفاده قرار بگیرد. در ابتدا تقسیم بندی داده‌های آموزشی و آزمایشی به صورت تصادفی انجام شد. ۸۰٪ از کلاس‌های هر مجموعه داده برای آموزش و ۲۰٪ از کلاس‌های هر مجموعه داده نیز برای آزمایش انتخاب گردید. به طور مثال در دیتاست AWA2 از ۴۰ کلاس برای آموزش و آزمایش کلاس‌های دیده شده استفاده شده و ۱۰ کلاس برای آزمایش کلاس‌های دیده نشده مورد استفاده قرار گرفته‌اند. در روش پیشنهادی مجموعه داده‌های دیده شده به دو بخش آموزش و آزمایش تقسیم شده است. در حالی که مجموعه داده‌های دیده نشده تنها شامل یک مجموعه آزمایشی است.

۴-۴- ساز و کار اصلاح

با توجه به اینکه دقت طبقه بندی تصاویر در یادگیری بدون نمونه تعمیم‌یافته اهمیت بسیار زیادی دارد در روش پیشنهادی سعی بر این است تا با کمک مدل‌هایی که دقت بالایی در این زمینه دارند ساز و کاری برای افزایش دقت طبقه بندی ارائه شود. بنابراین در روش پیشنهادی دو مدل با دقت طبقه بندی بالا انتخاب شده اند که به طور جداگانه فرآیند آموزش، استخراج ویژگی‌ها و طبقه بندی نمونه‌ها را انجام می‌دهند. طی طبقه بندی ممکن است سناریوهای مختلفی رخ دهد که با کمک ساز و کار اصلاح این سناریوها مدیریت می‌شوند. در شکل ۶ روند نما ساز و کار اصلاح نشان داده شده است.



شکل ۶: روند نمای ساز و کار اصلاح

در صورتی که طبقه بندی هر دو روش یکسان باشد نمونه در کلاس پیش بینی شده قرار می‌گیرد. در صورتی که طبقه بندی دو روش متفاوت باشد، ساز و کار اصلاح بررسی می‌نماید که آیا در هر دو کلاس پیش بینی شده حداقل یک نمونه وجود داشته باشد. سپس نمونه با یک نمونه تصادفی از هر کلاس مقایسه و در کلاسی که بیشترین شباهت را داشته قرار می‌گیرد. همچنین در صورت عدم وجود حداقل یک نمونه در دو کلاس پیش بینی شده، ساز و کار نمونه را به طور موقت طبقه بندی نمونه و اطلاعاتی چون نمونه و دو کلاس پیش بینی شده را ذخیره می‌نماید. سپس هر زمانی که حداقل یک نمونه در هر دو کلاس وجود داشت با مقایسه شباهت، نمونه را در کلاسی که بیشترین شباهت را داشته قرار می‌دهد.

۵- نتایج و بحث

برای ارزیابی و مقایسه دقت روش پیشنهادی یادگیری بدون نمونه تعمیم‌یافته ترکیبی با سایر روش‌ها باید شرایط به صورت یکسان فراهم شود. در این پژوهش برای ارزیابی دقت طبقه بندی تصاویر با یادگیری بدون نمونه تعمیم‌یافته مجموعه داده‌ها به دو بخش تقسیم شد تا بخشی از آن برای آموزش مدل و بخش دیگر برای آزمایش مدل مورد استفاده قرار بگیرد. با این کار بخشی از کلاس‌ها در فرآیند آموزش استفاده شده و با عنوان کلاس‌های دیده شده ارزیابی می‌شود. بخش دیگری از کلاس‌ها نیز که در بخش آموزش شرکت نداشته و تنها در بخش آزمایش مورد استفاده قرار می‌گیرند به عنوان کلاس‌های دیده نشده ارزیابی می‌شود.

۵-۱- تنظیمات پیاده سازی

برای پیاده سازی از زبان برنامه نویسی پایتون استفاده شد. در ابتدا هر کدام از روش‌های IDN-GZSL و DEGG پیاده سازی شد تا هر کدام جداگانه پس از پیش پردازش داده‌ها کار طبقه بندی را انجام دهند. در ابتدا این دو روش با کتابخانه‌های ResNet50 و VGG16 ویژگی‌های تصاویر را استخراج نموده و در ادامه نیز برچسب‌های نمونه‌ها ذخیره می‌شود. برای هر کلاس با کمک مدل آماری GMM اطلاعاتی چون پراکندگی و میانگین محاسبه شده تا در فرآیند داده‌های مصنوعی مورد استفاده قرار بگیرد. کار طبقه بندی نمونه‌ها توسط دو تابع به طور همزمان انجام شده و در نهایت ساز و کار اصلاح برای مقایسه طبقه بندی انجام می‌شود. با توجه به اینکه برای این پژوهش از ۵ مجموعه داده استفاده گردید، حجم فایل‌ها حدود ۲۰ گیگابایت بوده و شامل تعداد زیادی تصاویر می‌باشد. اجرای کد بر روی سیستم با مشخصات پردازنده Intel(R) Core i5-4210U و حافظه رم ۶ GB انجام گردید. داده‌های آموزشی و آزمایشی با نسبت ۴/۵ و ۱/۵ علامت گذاری شد تا پس از اجرای کد بتوان دقت

کلاس‌های دیده شده و دیده نشده را برآورد نمود. لازم به ذکر است کلاس‌های آموزشی به صورت تصادفی انتخاب گردید. همچنین با توجه به اینکه همزمان کار آموزش و طبقه بندی توسط دو روش مختلف انجام می‌گیرد زمان زیادی برای اجرای کد بر روی مجموعه داده‌ها لازم است.

۵-۲- معیار ارزیابی

ارزیابی در جدولی با مشخص کردن دقت در کلاس‌های دیده شده ، دقت در کلاس‌های دیده نشده و میانگین هارمونیک ارائه شده و با دقت روش‌های قبلی مقایسه می‌گردد. در طبقه بندی تصاویر با یادگیری بدون نمونه تعمیم‌یافته باید دقت در کلاس‌های دیده شده (S) و دیده نشده (U) به تفکیک محاسبه شود. دقت یا precision بیانگر آن است که چه میزان از نتایج بدست آمده صحیح می‌باشد و با فرمول (۱) محاسبه می‌شود :

$$\text{Precision} = TP / (TP+FP) \qquad (۱)$$

همچنین بازیابی یا Recall که به آن حساسیت نیز گفته می‌شود ، نشان دهنده میزان درصدی است که روش درست تشخیص داده است و با فرمول (۲) محاسبه می‌شود :

$$\text{Recall} = TP / (TP+FN) \qquad (۲)$$

سپس برای محاسبه میانگین دقت کلاس‌های دیده شده و دیده نشده از میانگین هارمونیک (H) با فرمول (۳) استفاده می‌شود :

$$H = (2*U*S)/U+S \qquad (۳)$$

۵-۳- ارزیابی روش روی دیتاست CUB

اولین مجموعه داده‌ای که برای ارزیابی مدل انتخاب گردید CUB است که دارای ۲۰۰ کلاس می‌باشد. برای آموزش مدل از ۱۶۰ کلاس استفاده شد. در نهایت کل دیتاست مورد ارزیابی قرار گرفته تا دقت کلاس‌های دیده شده و کلاس‌های دیده نشده مشخص گردد.

$$\text{دقت کلاس‌های دیده شده (S)} = \frac{۷۵۸۳}{۹۴۱۲} = ۸۰.۵۷٪$$

$$\text{دقت کلاس‌های دیده نشده (U)} = \frac{۱۷۰۱}{۲۳۷۹} = ۷۱.۵۰٪$$

$$\text{میانگین هارمونیک (H)} = \frac{S \times U \times ۲}{U + S} = ۷۵.۷۶٪$$

۵-۴- ارزیابی روش روی دیتاست SUN

دیتاست SUN یکی از مجموعه داده‌هایی می‌باشد که دارای بیشترین تعداد کلاس می‌باشد و ۷۱۷ دسته بوده که برخی دارای زیر کلاس‌هایی می‌باشد و در مجموع دارای ۸۹۹ کلاس است. برای بخش آموزش ۵۶۹ دسته انتخاب گردید که برای ارزیابی کلاس‌های دیده شده مورد استفاده قرار می‌گیرند و ۱۴۸ کلاس نیز به عنوان کلاس‌های دیده نشده انتخاب شد که تنها در زمان آزمایش استفاده شدند. در این پژوهش از کلاس "misc" صرف نظر شده است.

$$\text{دقت کلاس‌های دیده شده (S)} = \frac{۶۹۰۴}{۱۲۳۸۱} = ۵۵.۷۶٪$$

$$\text{دقت کلاس‌های دیده نشده (U)} = \frac{۱۳۵۲}{۲۶۰۷} = ۵۱.۸۶٪$$

$$\text{میانگین هارمونیک (H)} = \frac{S \times U \times ۲}{U + S} = ۵۳.۷۴٪$$

۵-۵- ارزیابی روش روی دیتاست FLO

مجموعه داده FLO نیز دارای ۱۰۲ کلاس می‌باشد که ۸۱ کلاس در آموزش مورد استفاده قرار گرفت و کلاس‌های باقیمانده نیز در بخش استفاده شدند.

$$\text{دقت کلاس‌های دیده شده (S)} = \frac{۵۲۵۶}{۱۶۳۲۱} = ۸۳.۱۵٪$$

$$\text{دقت کلاس‌های دیده نشده (U)} = \frac{۱۴۳۷}{۱۸۶۸} = ۷۶.۹۳٪$$

$$\text{میانگین هارمونیک (H)} = \frac{S \times U \times ۲}{U + S} = ۷۹.۹۲٪$$

۵-۶- ارزیابی روش روی دیتاست AWA2

روش ارائه شده بر روی دیتاست AWA2 نیز انجام شد. این مجموعه داده دارای ۵۰ کلاس از حیوانات بوده که برای بخش آموزش ۴۰ کلاس انتخاب گردید. در نهایت برای برآورد دقت روی کلاس‌های دیده شده و دیده نشده کلاس‌ها آزمایش شدند.

$$\text{دقت کلاس‌های دیده شده (S)} = \frac{۲۷۱۱۵}{۳۰۶۰۰} = ۸۸.۶۱٪$$

$$\text{دقت کلاس‌های دیده نشده (U)} = \frac{۴۲۴۴}{۶۷۲۲} = ۶۳.۱۴٪$$

$$\text{میانگین هارمونیک (H)} = \frac{S \times U \times ۲}{U + S} = ۷۳.۷۴٪$$

۵-۷- ارزیابی روش روی دیتاست aPY

آخرین مجموعه داده‌ای که برای ارزیابی روش پیشنهادی مورد استفاده قرار گرفته دیتاست aPY است که دارای ۳۲ کلاس بوده و ۲۵ کلاس برای بخش آموزش مورد استفاده قرار گرفت.

$$\text{دقت کلاس‌های دیده شده (S)} = \frac{۴۰۷۲}{۵۲۲۵} = ۷۷.۹۳٪$$

$$\text{دقت کلاس‌های دیده نشده (U)} = \frac{۱۳۹۶}{۳۴۲۲} = ۴۰.۷۹٪$$

$$\text{میانگین هارمونیک (H)} = \frac{S \times U \times ۲}{U + S} = ۵۳.۵۵٪$$

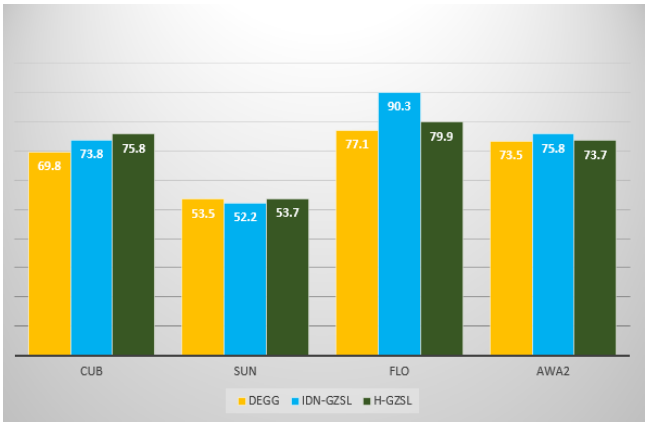
جدول ۱: مقایسه دقت روش پیشنهادی H-GZSL با روش‌های دیگر

Method	CUB			SUN			FLO			AWA2			aPY		
	U	S	H	U	S	H	U	S	H	U	S	H	U	S	H
JGM-ZSL [۴۰]	۴۲.۷	۴۵.۶	۴۴.۱	۴۴.۴	۳۰.۹	۳۶.۵	-	-	-	۵۶.۲	۷۱.۷	۶۳.۰	۳۱.۳	۴۳.۳	۳۶.۲
MSDN [۴۱]	۶۸.۷	۶۷.۵	۶۸.۱	۵۲.۲	۳۴.۲	۴۱.۳	-	-	-	۶۲.۰	۷۴.۵	۶۷.۷	-	-	-
CE-GZSL [۴۲]	۶۳.۹	۶۶.۸	۶۵.۳	۴۸.۸	۳۸.۶	۴۳.۱	۶۹.۰	۷۸.۷	۷۳.۵	۶۳.۱	۷۸.۶	۷۰.۷	-	-	-
DFCA-GZSL [۳۰]	۷۰.۹	۶۳.۱	۶۶.۸	۴۸.۹	۳۸.۸	۴۳.۳	۸۸.۰	۹۰.۳	۸۹.۲	۶۶.۵	۸۱.۵	۷۳.۳	-	-	-
VSCM-GZSL [۴۳]	۶۴.۶	۶۶.۸	۶۵.۷	۴۷.۰	۳۷.۶	۴۱.۸	۸۵.۳	۹۲.۱	۸۸.۶	۶۸.۶	۷۹.۰	۷۳.۴	-	-	-
CRnet [۴۴]	۴۵.۵	۵۶.۸	۵۰.۵	۳۴.۱	۳۶.۵	۳۵.۳	-	-	-	۵۲.۶	۷۸.۸	۶۳.۱	۳۲.۴	۶۸.۴	۴۴.۰
E-PGN [۴۵]	۵۲.۰	۶۱.۱	۵۶.۲	-	-	-	۷۱.۵	۸۲.۲	۷۶.۵	۵۲.۶	۸۳.۵	۶۴.۶	-	-	-
IDN-GZSL [۳۷]	**۷۷.۶**	۷۰.۲	۷۳.۸	**۶۰.۰**	۴۶.۲	۵۲.۲	**۹۰.۲**	**۹۰.۴**	**۹۰.۳**	**۶۶.۷**	۸۷.۷	**۷۵.۸**	-	-	-
DEGG [۳۵]	۶۰.۶	**۸۲.۳**	۶۹.۸	۴۹.۵	**۵۸.۲**	۵۳.۵	۷۲.۷	۸۲.۱	۷۷.۱	۶۰.۱	**۸۹.۹**	۷۳.۵	**۴۲.۱**	۷۷.۷	**۵۴.۶**
H-GZSL *	۷۱.۵	۸۰.۶	**۷۵.۸**	۵۱.۹	۵۵.۷	**۵۳.۷**	۷۶.۹	۸۳.۱	۷۹.۹	۶۳.۱	۸۸.۶	۷۳.۷	۴۰.۸	**۷۷.۹**	۵۳.۶

U: کلاس های دیده شده **S**: کلاس های دیده شده **H**: میانگین هارمونیک

۵-۸- ارزیابی کلی روش پیشنهادی

با اجرای مدل پیشنهادی بر روی ۵ مجموعه داده مد نظر دقت این روش در کلاس‌های دیده شده، دیده نشده و هارمونیک هر دیتاست مشخص شده که در جدول زیر قرار گرفته است. همچنین برای مقایسه بهتر با روش‌های انتخاب شده و روش‌های قبلی در زمینه یادگیری بدون نمونه تعمیم‌یافته، دقت روش‌های فوق نیز در جدول ۱ آورده شده است. یکی از نکات مهمی که با مقایسه روش‌های یادگیری بدون نمونه تعمیم‌یافته مشاهده می‌شود، دقت متفاوت روش‌ها در کلاس‌های دیده شده و دیده نشده می‌باشد. به همین دلیل برخی از روش‌ها در شناسایی و طبقه بندی کلاس‌های دیده شده و برخی دیگر در شناسایی و طبقه بندی کلاس‌های دیده نشده بهتر عمل می‌کنند. در شکل ۷ مقایسه بین دقت روش پیشنهادی با روش‌های دیگر نمایش داده شده است.

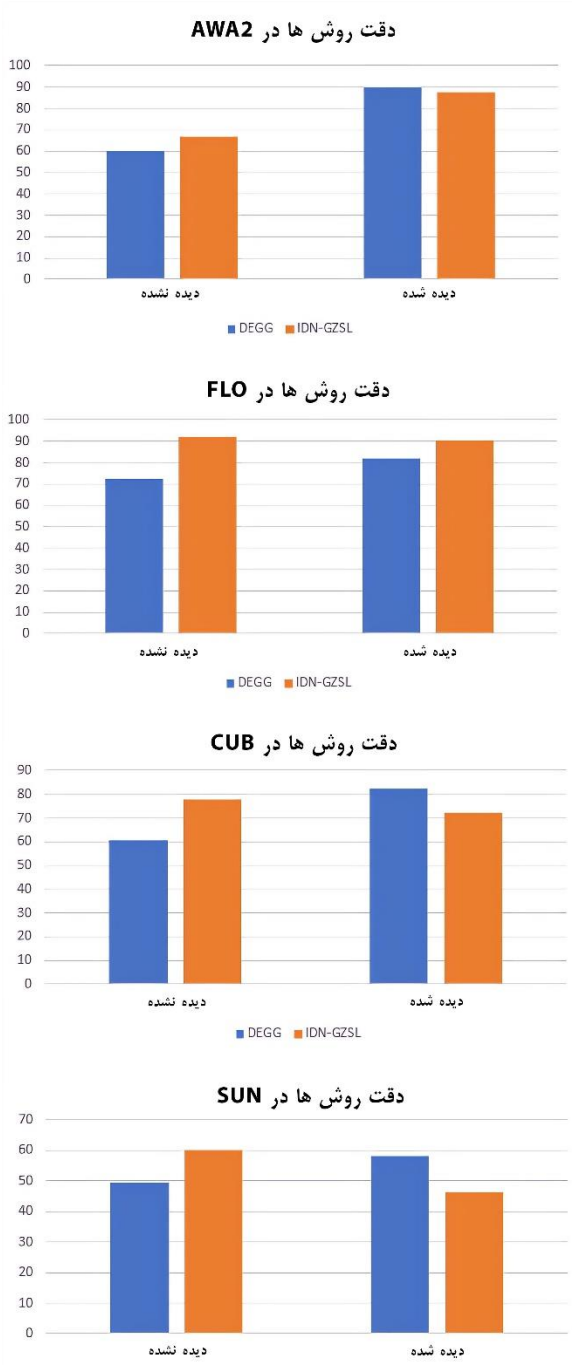


شکل ۷: نمودار مقایسه میانگین هارمونیک روش پیشنهادی H-GZSL با روش‌های دیگر بر روی ۴ مجموعه داده

۵-۹- بحث

با توجه به مرتبط بودن و دقت بالای روش‌های DEGG و IDN-GZSL، مقایسه‌ای بین این دو روش و دقت کلاس‌های دیده شده و دیده نشده انجام دادیم. در شکل ۸ مشاهده می‌شود که روش DEGG دقت بالاتری را در کلاس‌های دیده شده دارد. از سوی دیگر روش IDNL-GZSL نیز دقت بهتری را در زمینه تشخیص و طبقه بندی کلاس‌های دیده نشده دارد. روش یادگیری بدون نمونه تعمیم‌یافته ترکیبی با ترکیب این دو روش سعی کرده تا در هر زمینه با ساز و کار اصلاح تعادل نسبی را بین طبقه بندی دو روش داشته باشد. با آنکه روش پیشنهادی نتوانسته در بحث کلاس‌های دیده شده و دیده نشده به دقت بیشتری دست پیدا نکند، اما با توجه به تعادلی که در هر مورد بین دو روش ایجاد نموده سبب شده تا دقت میانگین در هر بخش افزایش پیدا کند. با افزایش دقت در کلاس‌های دیده شده و دیده نشده، دقت کلی یا میانگین هرمونیک روش پیشنهادی نیز افزایش پیدا نموده است.

از این رو روش یادگیری بدون نمونه تعمیم‌یافته ترکیبی توانسته در مجموعه داده‌های CUB و SUN دقت بالاتری را بدست بیاورد. همچنین در مجموعه داده‌های FLO و AWA2 با آنکه روش پیشنهادی توانسته دقت مناسبی را داشته باشد اما به دلیل نزدیک بودن ویژگی‌های بصری کلاس‌ها، ساز و کار اصلاح نتوانسته با دقت بالاتری نمونه‌ها را در طبقه بندی مناسب قرار دهد. در شکل ۹ به خوبی مشاهده می‌شود که روش پیشنهادی H-GZSL توانسته تعادل مناسبی را بین روش‌ها در کلاس‌های دیده شده و دیده نشده به وجود بیاورد.

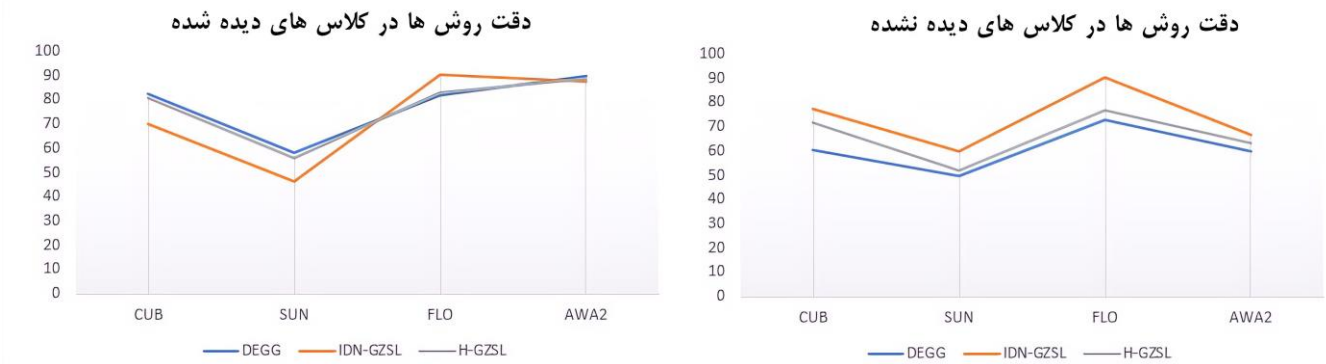


شکل ۸: مقایسه دقت روش‌ها در کلاس‌های دیده شده و دیده نشده

یکی از مواردی که با پیاده سازی و اجرای کد و در نهایت تحلیل نتایج مشخص گردید نحوه انتخاب داده‌ها می‌باشد. انتخاب اینکه چه مقدار از داده‌ها در بخش آموزش و چه مقدار در بخش آزمایش استفاده شوند بر روی دقت تاثیر مستقیمی دارد. از طرف دیگر تعداد نمونه‌های کلاس‌ها نیز متفاوت بوده و انتخاب کلاس‌هایی با تعداد نمونه‌های بیشتر یا کمتر نیز می‌تواند باعث تغییر دقت بدست آمده شود. مجموعه داده‌های CUB و SUN به سبب داشتن تعداد نمونه‌ها و کلاس‌های بسیار می‌توانند به خوبی مدل‌های پیشنهادی را به چالش بکشند. روش پیشنهادی H-GZSL توانسته در مقایسه با روش‌های دیگر دقت تشخیص بالایی را در این دو دیتاست بدست بیاورد. همچنین مجموعه داده SUN به دلیل داشتن ۸۹۹ کلاس مختلف نزدیک به فضای واقعی بوده و باعث شده تا به خوبی عملکرد روش‌های ارائه شده در شرایط واقعی مشخص شود. روش پیشنهادی H-GZSL دقت تشخیص بالایی در مقایسه با روش‌های قبلی دارا می‌باشد. روش E-PGN بر روی این مجموعه داده اجرا نشده است.

۶- نتیجه‌گیری

الگوریتم‌های یادگیری ماشین معمولاً برای بهبود عملکرد خود به مقادیر زیادی از داده‌های آموزشی نیاز دارند. با این حال در بسیاری از موارد دسترسی به داده‌های بزرگ و متنوع ممکن نیست که این امر به چالش‌های جدیدی منجر می‌شود. یادگیری بدون نمونه به شرایطی اشاره دارد که مدل باید بتواند بدون داشتن هیچ نمونه‌ای از کلاس هدف در داده‌های آموزشی عملکرد مناسبی را ارائه دهد. یادگیری بدون نمونه تعمیم‌یافته یک گام فراتر از یادگیری بدون نمونه است. در این نوع یادگیری مدل نه تنها باید کلاس‌های دیده نشده را شناسایی کند بلکه باید بتواند بین کلاس‌های دیده شده و ندیده شده تمایز قائل شود و به درستی آن‌ها را طبقه بندی کند. یکی از مهم ترین ویژگی‌های یادگیری بدون نمونه تعمیم‌یافته نیاز به تعداد کمتری داده‌های آموزشی می‌باشد. با ارائه یادگیری بدون نمونه تعمیم‌یافته ترکیبی امکان طبقه بندی نمونه‌ها را با دو روش



شکل ۹: مقایسه روش پیشنهادی در کلاس‌های دیده شده و دیده نشده

با رویکردهای مختلف را فراهم نمودیم. این تفاوت رویکرد باعث شده تا هر کدام از روش‌ها در طبقه بندی کلاس‌های دیده شده و دیده نشده بهتر عمل نمایند.

ترکیب این دو روش و معرفی ساز و کار اصلاح نیز باعث شده تا ضمن مقایسه طبقه بندی‌ها، در زمانی که روش‌ها در طبقه بندی تناقض و اختلاف دارند توسط این ساز و کار نمونه‌ها مقایسه شده و در مناسب ترین دسته قرار بگیرند. روش پیشنهادی تعادل نسبی بین دو روش در طبقه بندی‌های کلاس‌های دیده شده و دیده نشده برقرار کرده است. از این رو در مجموعه داده‌هایی که تنوع بیشتری دارند این روش عملکرد بهتری دارد. روش یادگیری بدون نمونه تعمیم‌یافته ترکیبی توانسته دقت بالاتری را در میانگین هارمونیک در مقایسه با روش‌های پیشین در مجموعه داده‌های CUB و SUN داشته باشد.

۷- کارهای آینده

طی مرور پژوهش‌ها و مقالات قبلی مشاهده گردید که روش‌های یادگیری بدون نمونه تعمیم‌یافته جدا از دقت کلی که دارا می‌باشند در مورد طبقه بندی کلاس‌های دیده شده و دیده نشده نیز دقت‌های متفاوتی را دارند. برخی از روش‌ها طبقه بندی کلاس‌های دیده شده را بهتر انجام داده و برخی دیگر نیز طبقه بندی کلاس‌های دیده نشده را با دقت بالاتری انجام می‌دهند. با توجه به اینکه هنوز هم فضا برای پژوهش و افزایش دقت در این زمینه وجود دارد می‌توان از یادگیری بدون نمونه تعمیم‌یافته ترکیبی برای بهبود عملکرد همزمان استفاده نمود. از این رو می‌توان برای کارهای آینده بر روی بهبود ساز و کار اصلاح برای تعادل بهتر مقایسه بین دو روش کار شود. در برخی مجموعه داده‌ها کلاس‌هایی با یک نمونه زیاد بوده و همین بر روی بخش مقایسه‌ای ساز و کار اصلاح اثر گذار بوده که می‌توان این شرایط را نیز در ساز و کار تعریف نمود.

مراجع

[1] X. Wang, A. Liu, L. Wu, C. Li, Y. Liu and X. Chen, "A Generalized Zero-Shot Learning Scheme for SSVEP-Based BCI System," in IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 31, pp. 863-874, 2023.

[2] L. Niu, J. Cai, A. Veeraraghavan, and L. Zhang, "Zero-shot learning via category-specific visual-semantic mapping and label refinement", IEEE Transactions on Image Processing, vol. 28, no. 2, pp. 965–979, 2018.

[3]] H. Jiang, R. Wang, S. Shan, and X. Chen, "Learning class prototypes via structure alignment for zero-shot recognition," in Proceedings of the European Conference on Computer Vision, 2018, pp. 121–138.

[4] Wei Wang, "Generalized Zero-Shot Activity Recognition with Embedding-Based Method", ACM Transactions on Sensor Networks, 19, 1–25, , 2023.

[5] Qian Wang, "Generalized zero-shot domain adaptation via coupled conditional variational autoencoders", Neural Networks, 163, 40-52, 2023.

[6] E. Schonfeld, S. Ebrahimi, S. Sinha, T. Darrell, and Z. Akata, "Generalized zero-and few-shot learning via aligned variational autoencoders," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019, pp. 8247–8255.

[7] Y. Ye, Y. He, T. Pan, J. Li and H. T. Shen, "Alleviating Domain Shift via Discriminative Learning for Generalized Zero-Shot Learning, " in IEEE Transactions on Multimedia, vol. 24, pp. 1325-1337, 2022.

[8] Y. An, S. Kim, Y. Liang, R. Zimmermann, D. Kim and J. Kim, "Content-Attribute Disentanglement for Generalized Zero-Shot Learning, " in IEEE Access, vol. 10, pp. 58320-58331, 2022.

[9] A. Paul, N. C. Krishnan, and P. Munjal, "Semantically aligned bias reducing zero shot learning," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019, pp. 7056–7065.

[10] Richard Socher, "zero-shot learning through cross-modal transfer", 2013.

[11] Wei-Lun Chao, "an empirical study and analysis of generalized zero-shot learning for object recognition in the wild", 2017.

[12] Maxime Bucher, "Generating Visual Representations for Zero-Shot Classification", 2017 IEEE International Conference on Computer Vision Workshops (ICCVW), 2017.

[13] Kibok Lee, "hierarchical novelty detection for visual object recognition", 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018.

[14] Yongqin Xian, "Feature Generating Networks for Zero-Shot Learning", 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018.

[15] Haofeng Zhang, "Triple Verification Network for Generalised Zero-shot Learning", IEEE Transactions on Image Processing, 2019, 28, 506 – 517.

[16] Yashas Annadani, "Preserving Semantic Relations for Zero-Shot Learning", 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018.

[17] Pengkai Zhu, "Generalized Zero-Shot Recognition based on Visually Semantic Embedding", 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019.

[18] Vinay Kumar Verma, "Towards Zero-Shot Learning with Fewer Seen Class Examples", 2021 IEEE Winter Conference on Applications of Computer Vision (WACV), 2020.

[19] Lu Liu, "Attribute Propagation Network for Graph Zero-shot Learning", AAAI Conference on Artificial Intelligence, 2020.

[20] Jie Shao, "Generalized Zero-Shot Learning With Multi-Channel Gaussian Mixture VAE", IEEE Signal Processing Letters, 2020, 27, 456 – 460.

[21] Xiao Li, Min Fang, "Learning domain invariant unseen features for generalized zero-shot classification", Knowledge-Based Systems, Volume 206, 2020.

[22] Tingting Xu, "Dual Generative Network with Discriminative Information for Generalized Zero-Shot Learning", Complexity, 2021.

[23] Zhi Chen, "Entropy-Based Uncertainty Calibration for Generalized Zero-Shot Learning", Databases Theory and Applications, 2021.

[24] Yang, Shiqi & Wang, Kai, "On Implicit Attribute Localization for Generalized Zero-Shot Learning", IEEE Signal Processing Letters, PP. 1-1. 10.1109/LSP, 2021.

[25] J. Ding, X. Hu and X. Zhong, "A Semantic Encoding Out-of-Distribution Classifier for Generalized Zero-Shot Learning," in IEEE Signal Processing Letters, vol. 28, pp. 1395-1399, 2021.

[26] L. Feng and C. Zhao, "Transfer Increment for Generalized Zero-Shot Learning," in IEEE Transactions on Neural Networks and Learning Systems, vol. 32, no. 6, pp. 2506-2520, June 2021.

[27] X. Chen, J. Li, X. Lan and N. Zheng, "Generalized Zero-Shot Learning Via Multi-Modal Aggregated Posterior Aligning Neural Network," in IEEE Transactions on Multimedia, vol. 24, pp. 177-187, 2022.

[28] T. Shermin, S. W. Teng, F. Sohel, M. Murshed and G. Lu, "Bidirectional Mapping Coupled GAN for Generalized Zero-Shot Learning," in IEEE Transactions on Image Processing, vol. 31, pp. 721-733, 2022.

[29] G. -S. Xie et al., "Generalized Zero-Shot Learning With Multiple Graph Adaptive Generative Networks," in IEEE Transactions on Neural Networks and Learning Systems, vol. 33, no. 7, pp. 2903-2915, July 2022.

[30]] H. Su, J. Li, K. Lu, L. Zhu and H. T. Shen, "Dual-Aligned Feature Confusion Alleviation for Generalized Zero-Shot Learning," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 33, no. 8, pp. 3774-3785, Aug. 2023.

[31] Z. Wang, Y. Gou, J. Li, L. Zhu and H. T. Shen, "Language-Augmented Pixel Embedding for Generalized Zero-Shot Learning," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 33, no. 3, pp. 1019-1030, March 2023.

[32] Y. Yang, X. Zhang, M. Yang and C. Deng, "Adaptive Bias-Aware Feature Generation for Generalized Zero-Shot Learning," in IEEE Transactions on Multimedia, vol. 25, pp. 280-290, 2023.

[33] Y. Zhang, S. Feng, "Enhancing Domain-Invariant Parts for Generalized Zero-Shot Learning," in Proceedings of the 31st ACM International Conference on Multimedia, pp. 6283–6291, 2023.

[34] Y. Ye, T. Pan, T. Luo, J. Li and H. T. Shen, "Learning MLatent Representations for Generalized Zero-Shot Learning," in IEEE Transactions on Multimedia, vol. 25, pp. 2252-2265, 2023.

[35] J. Cavazza, V. Murino and A. D. Bue, "No Adversaries to Zero-Shot Learning: Distilling an Ensemble of Gaussian Feature Generators," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 45, no. 10, pp. 12167-12178, Oct. 2023.

[36] J. Wu, T. Zhang, Z. -J. Zha, J. Luo, Y. Zhang and F. Wu, "Prototype-Augmented Self-Supervised Generative Network for Generalized Zero-Shot Learning, " in IEEE Transactions on Image Processing, vol. 33, pp. 1938-1951, 2024.

[37] Liang, Y., Cao, W. Isolation and distillation network for generalized zero-shot learning. Neural Comput & Applic 36, 13935–13955, 2024.

[38] Farhad Pourpanah, "A Review of Generalized Zero-Shot Learning Methods, IEEE Transactions on Pattern Analysis and Machine Intelligence, 45, 4051 – 4070, 2023.

[39] Li, M. Ye, L. Zhou, D. Zhang, C. Zhu, and Y. Liu, "Improving generalized zero-shot learning by semantic discriminator," arXiv preprint arXiv:2005.13956, 2020.

[40] R. Gao et al., "Zero-VAE-GAN: Generating Unseen Features for Generalized and Transductive Zero-Shot Learning," in IEEE Transactions on Image Processing, vol. 29, pp. 3665-3680, 2020.

[41] Chen, Shiming, Hong, Ziming, Xie, Guo-Sen, Yang, Wenhan, Peng, Qinmu, Wang, Kai, Zhao, Jian and You, Xinge. "MSDN: Mutually Semantic Distillation Network for Zero-Shot Learning", Paper presented at the meeting of the CVPR, 2022.

[42] Z. Han, Z. Fu, S. Chen and J. Yang, "Contrastive Embedding for Generalized Zero-Shot Learning," in 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 2021.

[43] Zhenqi Z, Cao W, "Visual-semantic consistency matching network for generalized zero-shot learning", Neurocomputing 536:30–39, 2023.

[44] F. Zhang and G. Shi, "Co-representation network for generalized zero-shot learning," in Proc. Int. Conf. Mach. Learn., pp. 7434–7443, 2019.

[45] Y. Yu, Z. Ji, J. Han, and Z. Zhang, "Episode-based prototype generating network for zero-shot learning," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2020, pp. 14032–14041.

پاورقی‌ها:

[1] zero-shot learning
[2] Generalized zero-shot learning
[3] Hubness
[4] Generative Adversarial Network
[5] Class-level Over-fitting
[6] Variational Autoencoder
[7] Attribute Propagation Network
[8] K-Nearest Neighbors
[9] Multi-channel Gaussian Mixture VAE
[10] Simple and Effective Localized Attribute Representation
[11] Global Max Pooling
[12] Global Average Pooling
[13] Out-Of-Distribution
[14] Local Relational Knowledge
[15] Global Relational Knowledge
[16] Wasserstein Autoencoder
[17] Dual-Aligned Feature Confusion Alleviation
[18] Conditional Invertible Neural Network
[19] Language-Augmented Pixel Embedding
[20] Enhancing doMain-invariant Parts
[21] Prototype-augmented Self-supervised Generative Network
[22] Isolation and Distillation Network