

طراحی و ارزیابی یک شبکه عصبی کپسولی جدید برای طبقه‌بندی نامتوازن تصاویر حامد جباری^۱ و نوشین بیگدلی^۲

چکیده

طبقه‌بندی نامتوازن تصاویر یکی از مسائل مهم و دشوار در زمینه داده‌کاوی است. با عدم توانایی الگوریتم‌های طبقه‌بندی استاندارد، شبکه‌های عصبی کپسولی با در نظر گرفتن ارتباطات فضایی ویژگی‌ها، در مقایسه با سایر شبکه‌های عمیق مثل شبکه‌های عصبی کانولوشنی بستر مناسبی را برای طراحی مدل‌های طبقه‌بندی نامتوازن فراهم می‌کنند. از طرف دیگر چندشاخگی در ترک‌های سطحی یکی از ناهنجاری‌ها و دسته‌های اقلیت موجود در سازه‌های بتنی است که تشخیص آن می‌تواند در نگهداری سازه‌های بتنی و مدیریت هزینه‌ها موثر باشد. به همین منظور در این مقاله یک معماری جدید بر اساس شبکه‌های عصبی کپسولی برای ارزیابی طبقه‌بندی نامتوازن تصاویر ترک‌های سطحی در سازه‌های بتنی معرفی شده است. بررسی و مقایسه شبکه پیشنهادی با شبکه‌های کانولوشنی در طبقه‌بندی متوازن و نامتوازن ترک‌های سطحی روی ۱۳۵۰۰ مجموعه تصاویر جمع‌آوری شده، نشان از برتری شبکه پیشنهادی داشت. شبکه پیشنهادی در بررسی اثر کاهش تعداد تصاویر آموزش در دقت طبقه‌بندی نیز برتری چشم‌گیری در مقایسه با شبکه‌های کانولوشنی از خود نشان داد. این شبکه طبقه‌بندی متوازن ترک‌های سطحی را با دقت ۹۹/۵۶ درصد انجام داد. هم‌چنین شبکه پیشنهادی تا عدم توازن دسته اقلیت به اکثریت ۱ به ۸، دقت بالای ۸۰ درصد داشت که نسبت به سایر روش‌ها بسیار مناسب است.

کلیدواژه‌ها

طبقه‌بندی تصاویر، طبقه‌بندی نامتوازن، ترک‌های سطحی، چندشاخگی، یادگیری عمیق، شبکه‌های عصبی کپسولی

۱ مقدمه

طبقه‌بندی مجموعه‌ای از تصاویر در دسته‌ها یا دسته‌های مختلف باید رابطه بین تصاویر و دسته‌هایی که در آنها طبقه‌بندی شده‌اند، به خوبی درک شود. برای دستیابی به این مهم، سیستم طبقه‌بندی باید آموزش ببیند. آموزش، معمولاً با استفاده از تکنیک‌های یادگیری (یادگیری با نظارت و یادگیری بدون نظارت) و مجموعه‌های ویژگی صورت می‌گیرد [۱]. طبقه‌بندی تصاویر معمولاً با چالش‌هایی مثل تنوع درون دسته‌ای، تنوع مقیاس تصاویر، تنوع دید، انسداد، شدت روشنایی تصاویر و بی‌نظمی در پس‌زمینه همراه است [۲]. از طرفی، اغلب مجموعه تصاویر

امروزه با رشد تولید تصاویر دیجیتال در حوزه‌های مختلف از قبیل راه و ساختمان، شبکه‌های اجتماعی، پزشکی، نانوذرات، پتروشیمی، تشخیص عیب و ... نیاز به طبقه‌بندی تصاویر بیش از گذشته احساس می‌شود. طبقه‌بندی تصاویر فرآیند اختصاص برجسب به تصاویر با توجه به نوع آنها (دسته‌ها) است. برای این مقاله در بهمن‌ماه سال ۱۳۹۹ دریافت، در فروردین‌ماه سال ۱۴۰۰ بازنگری و در اردیبهشت‌ماه همان سال پذیرفته شد.

^۱ دانشجوی دکتری مهندسی برق - کنترل، دانشگاه بین‌المللی امام خمینی (ع) قزوین.

رایانامه: h_jabbari@edu.ikiu.ac.ir

^۲ گروه مهندسی برق - کنترل، دانشگاه بین‌المللی امام خمینی (ع) قزوین.

رایانامه: n.bigdeli@eng.ikiu.ac.ir

نویسنده مسئول: نوشین بیگدلی

DOR: 20.1001.1.23831197.1401.9.1.1.2

¹ Viewpoint Variation

² Occlusion

رویداد ۲ درصد است، یک مدل طبقه‌بندی با دقت ۹۸ درصد، اگر همه نمونه‌ها را در دسته اکثریت طبقه‌بندی کند و مشاهدات مربوط به دسته اقلیت ۲ درصدی را به‌عنوان داده نویز حذف کند، مدل دقیقی نخواهد بود [۱۳].

۳- اختلالات کوچک و کمبود نمونه‌های دسته اقلیت، چالش‌هایی برای یادگیری نامتوازن است که اغلب باعث می‌شود مدل‌های یادگیری در تشخیص الگوهای نادر موفق نشوند [۱۱، ۱۴].

در دهه گذشته، بسیاری از روش‌های یادگیری ماشینی برای مقابله با طبقه‌بندی نامتوازن تصاویر توسعه یافته‌اند که بیشتر آن‌ها بر اساس تکنیک‌های پیش‌پردازش، یادگیری حساس به هزینه و یادگیری عمیق است [۱۲، ۱۵]. تکنیک‌های پیش‌پردازش اغلب قبل از ساخت مدل یادگیری به‌منظور دستیابی به اطلاعات ورودی بهتر انجام می‌شود. در واقع، در این رویکرد، برای حل عدم تعادل، توزیع دسته‌ها نیز دستکاری شده تا تعادل در کل دسته‌های مجموعه داده‌ها حاصل شود. رویکردهای پیش‌پردازش شامل روش‌های نمونه‌برداری مجدد^۱ (در فضای نمونه) و روش‌های انتخاب و استخراج ویژگی است که سعی در بهینه کردن فضای ویژگی دارد [۱۱]. تکنیک‌های نمونه‌برداری مجدد که با افزایش داده در دسته اقلیت یا کاهش داده در دسته اکثریت و یا با ترکیبی از این دو سعی در توازن بین دسته‌ها دارند، محبوبیت بالاتری نسبت به سایر تکنیک‌ها دارند. البته با تخصیص توابع هزینه مناسب به تکنیک‌های نمونه‌برداری مجدد می‌توان کارایی آن‌ها را بهتر هم کرد. در مقایسه با پیشرفت‌هایی که در روش‌های نمونه‌گیری مجدد حاصل شده است، مقالات کمتری انتخاب و استخراج ویژگی را در نظر گرفته‌اند. پیش‌تر ذکر گردید که نمونه‌های دسته اقلیت می‌توانند به راحتی به‌عنوان یک نویز شناخته و حذف شوند. با این حال می‌توان با انتخاب و استخراج ویژگی‌های مناسب، ویژگی‌های نامناسب را در فضای ویژگی حذف کرد و این خطر را کاهش داد [۱۶]. استراتژی دوم برای مقابله با مسائل طبقه‌بندی نامتوازن تصاویر، یادگیری حساس به هزینه است که در آن ماتریس‌های هزینه نقش کلیدی برای جلوگیری از طبقه‌بندی غلط نمونه‌های دسته اقلیت دارند. استراتژی سوم نیز استفاده از شبکه‌های عمیق نظیر شبکه‌های عصبی کانولوشنی و شبکه‌های عصبی کپسولی و یا با ترکیب این شبکه‌ها با استراتژی‌های قبلی است [۱۷]. جدول ۱ خلاصه‌ای از استراتژی‌های اصلی برای طبقه‌بندی نامتوازن تصاویر را نشان می‌دهد.

یکی از ناهنجاری‌های موجود در سازه‌های بتنی، چندشاخگی در ترک‌های سطحی می‌باشد که شناسایی آن در نظارت، نگهداری و ایمنی پل‌ها، سدها، جاده‌ها، تونل‌ها، آسمان‌خراش‌ها و ... موثر است. نقطه‌ای که ترک‌های سطحی در سازه‌های بتنی به دو یا چند ترک جدید تبدیل شود، چندشاخگی در ترک نامیده می‌شود [۱۰]. این چندشاخه شدن و تغییر زاویه (که معمولاً نسبت به سایر

دنیای واقعی، ترکیبی از نمونه‌های طبیعی و درصد کمی نمونه‌های غیرطبیعی یا جالب است و به‌نوعی نامتوازی در آن‌ها وجود دارد. تشخیص وقایع غیرطبیعی و یا یک ناهنجاری، یک مسئله پیش‌بینی یا به‌طور معمول یک مسئله طبقه‌بندی نامتوازن است که یک چالش مهم در طبقه‌بندی تصاویر می‌باشد.

شناسایی رفتارهای نادر، غیرطبیعی و غیرمعمول در تصاویر، کاری دشوار است و شناسایی چنین مواردی غالباً منجر به ایجاد چالش‌های فراوانی می‌شود. طبق تعریف، وقایع نادر به رویدادهایی اشاره دارند که بسیار کمتر از وقایع معمول رخ می‌دهند. وجود نمونه‌ها و یا تصاویر بسیار کم از وقایع نادر، باعث کاهش کارایی عمل تشخیص شده و آن را تبدیل به مسئله طبقه‌بندی نامتوازن تصاویر می‌کند. مجموعه نامتوازن تصاویر به مجموعه تصاویری اشاره دارد که در آن، یک یا چند دسته شامل نمونه‌های بسیار بیشتری نسبت به سایر دسته‌ها است. رایج‌ترین دسته، دسته اکثریت و کمیاب‌ترین دسته، دسته اقلیت نامیده می‌شود که در این میان، اکثریت دسته اقلیت مورد توجه است [۳]. بسیاری از موضوعات مهم در مباحث پزشکی نظیر تشخیص تومورهای بدخیم موجود در بافت پستان [۴، ۵]، تشخیص گلبول‌های سرطانی خون [۶]، شناسایی تومورهای مغز [۷] و تشخیص ناهنجاری‌های موجود در استخوان‌ها [۸]، نمونه‌هایی از مجموعه‌های نامتوازن هستند که تشخیص درست دسته اقلیت منجر به تشخیص درست ناهنجاری مربوطه می‌شود. تشخیص ترک و خوردگی موجود در لوله‌های خطوط انتقال در صنایع نفت و گاز [۹] و یا تشخیص ترک‌های سطحی در سازه‌های بتنی [۱۰] نیز از دیگر نمونه‌های مجموعه نامتوازن تصاویر است که در آنها، تشخیص دسته اقلیت از اهمیت بالایی برخوردار است. تشخیص وقایع نادر یا دسته اقلیت به دلیل کمیابی نمونه‌های آن‌ها دشوار است. با این حال، طبقه‌بندی اشتباه وقایع نادر ممکن است منجر به هزینه‌های سنگینی شود. به طور مثال در تشخیص بیماری سرطان، تعداد افراد مبتلا به سرطان بسیار کمتر از تعداد افراد سالم است؛ حال اگر فرد مبتلا به بیماری سرطان، سالم تشخیص داده شود، باعث تلف شدن زمان درمان و همچنین بروز فاجعه برای زندگی آن فرد می‌شود. بنابراین، طبقه‌بندی نامتوازن تصاویر به‌طور قابل توجهی با مدل‌های طبقه‌بندی سنتی در چالش است. این چالش‌ها از سه دلیل عمده ناشی می‌شوند:

۱- طبقه‌بندی‌های استاندارد مانند رگرسیون، ماشین بردار پشتیبانی و درخت تصمیم‌گیری، برای مجموعه‌های آموزشی متوازن، مناسب هستند. هنگام مواجهه با سناریوهای نامتوازن، اغلب این مدل‌ها نتایج طبقه‌بندی زیر بهینه ارائه می‌دهند [۱۱].

۲- فرآیند یادگیری با معیارهای عملکرد سراسری هدایت و سنجیده می‌شود [۱۲]. به‌طور مثال دقت پیش‌بینی موجب ایجاد تمایل نسبت به دسته اکثریت می‌شود؛ درحالی‌که اگر مدل پیش‌بینی نیز دقت بالایی داشته باشد، قسمت‌های نادر ناشناخته باقی می‌مانند. مثلاً وقتی هدف، کشف یک ناهنجاری با نرخ

¹ Resampling

تعمیر سازه‌های بتنی می‌شود [۲۵]. از این رو، تشخیص ناهنجاری مهمی مثل چندشاخگی در ترک‌های سطحی می‌تواند به جلوگیری از خسارت‌های احتمالی و حداقل رساندن هزینه‌ها کمک شایانی کند. مرسوم‌ترین روش برای بررسی ترک‌های سطحی، بازرسی بصری است که دارای معایبی همچون وقت‌گیر بودن، نداشتن قابلیت اطمینان بالا، عدم امکان استفاده در سازه‌های دارای ارتفاع است [۱۰]. به همین دلیل تکنیک‌های پردازش تصویر و یادگیری ماشین به‌عنوان یک روش جایگزین امیدوارکننده برای بررسی ترک‌های سطحی در سازه‌های بتنی معرفی شده‌اند [۲۶].

ترک‌های سطحی کمتر رخ می‌دهد) در مکانیک شکست الاستیک خطی خیلی مهم و قابل بررسی جهت وضعیت سازه‌های بتنی است. در واقع بالا بودن سطح شدت ترک‌های سطحی می‌تواند منجر به شکسته شدن و تبدیل آنها به چند ترک سطحی شود [۱۰]. بنابراین شناسایی چندشاخگی در ترک‌های سطحی به دلیل آنکه می‌تواند منجر به شناسایی سطوح شدت بالا شود، از اهمیت به سزایی در سازه‌های بتنی برخوردار است. همچنین با توجه به اینکه ترک‌های سطحی در اکثر سازه‌های بتنی یک معیار برای ماندگاری سازه به شمار می‌رود، هر ساله هزینه‌های زیادی صرف نگهداری و

جدول ۱ استراتژی‌های اصلی برای طبقه‌بندی نامتوازن تصاویر

روش‌های طبقه‌بندی نامتوازن تصاویر	روش		توضیحات
	تکنیک‌های پیش‌پردازش	بیش نمونه‌برداری	این روش با ایجاد نمونه‌های جدید برای گروه اقلیت سعی در از بین بردن چولگی در توزیع داده‌ها دارد. دو روش پرکاربرد برای ایجاد نمونه‌های اقلیت مصنوعی، تکثیر تصادفی نمونه‌های اقلیت و بیش-نمونه‌برداری مصنوعی دسته اقلیت هستند [۱۸].
زیر نمونه‌برداری		این روش با حذف نمونه‌های طبیعی در دسته اکثریت سعی در از بین بردن چولگی در توزیع داده‌ها دارد. ساده‌ترین و مؤثرترین روش، حذف تصادفی نمونه‌های دسته اکثریت است [۱۹].	
نمونه‌برداری ترکیبی		این روش ترکیبی از روش‌های بیش-نمونه‌برداری و زیر-نمونه‌برداری هستند و هم‌زمان با ایجاد نمونه‌های جدید در گروه اقلیت و حذف نمونه‌ها در گروه اکثریت سعی در ایجاد توازن بین داده‌ها دارد.	
انتخاب و استخراج ویژگی		هدف انتخاب ویژگی‌ها، انتخاب زیرمجموعه‌ای از k ویژگی از کل فضای ویژگی است که به طبقه‌بند اجازه می‌دهد به عملکرد مطلوب برسد، به طوری که k یک پارامتر انتخاب‌شده توسط کاربر یا یک پارامتر تطبیقی است. هدف از استخراج ویژگی، کاهش ابعاد است که تکنیک‌های مختلفی برای آن وجود دارد، مانند تجزیه و تحلیل مؤلفه اصلی، تجزیه مقادیر ویژه و فاکتورگیری ماتریس [۲۰].	
یادگیری حساس به هزینه		با اتخاذ هزینه‌های بالاتر برای طبقه‌بندی غلط نمونه‌های دسته اقلیت و با توجه به نمونه‌های دسته اکثریت، یادگیری حساس به هزینه می‌تواند در هر دو سطح داده (به‌عنوان مثال نمونه‌برداری مجدد و انتخاب ویژگی) و در سطح الگوریتمی گنجانده شود [۲۱]. هزینه‌ها اغلب به‌عنوان ماتریس‌های هزینه تعریف می‌شوند [۲۲]. به طوری که درایه‌های آن، هزینه‌های طبقه‌بندی اشتباه برای انتساب نمونه‌ها را نشان می‌دهد.	
	یادگیری عمیق	با استفاده از معماری‌های مختلف شبکه‌های عمیق و یا با ترکیبی از روش‌های گذشته سعی در طبقه‌بندی نامتوازن تصاویر دارند [۱۷، ۲۳، ۲۴].	

تصمیم‌گیری به منظور طبقه‌بندی نهایی تصاویر استفاده شد [۳۰]. این روش به طور متوسط ۸۰ درصد موفقیت در شناسایی نوع ترک را به دنبال داشت. همچنین دو الگوریتم هوشمند تشخیص ترک (تشخیص لبه و یادگیری عمیق) در مطالعه‌ای در سال ۲۰۱۸ با یکدیگر مقایسه شدند که به ترتیب صحتی معادل ۷۹/۹ و ۹۲/۰۸ درصد در پی داشتند [۳۱]. یک روش تشخیص ترک خودکار مبتنی بر بینایی ماشین با استفاده از شبکه عصبی کاملاً کانولوشنی با یک چارچوب رمزگذار و رمزگشا در سال ۲۰۱۹ پیشنهاد شد که تقریباً با دقت ۹۲ درصدی ترک‌های سطحی را طبقه‌بندی می‌کرد [۳۲]. طبقه‌بندی خودکار ترک‌های سطحی به دو دسته تصاویر دارای ترک و تصاویر بدون ترک با استفاده از شبکه عصبی کانولوشنی عمیق در سال ۲۰۲۰ انجام شد که منجر به درصد صحت بالای ۹۴ درصد شد [۳۳]. همچنین پژوهشی در همین سال با استفاده از شبکه‌های کانولوشنی، طبقه‌بندی ترک‌های سازه‌های بتنی را به دسته‌های بدون ترک، ترک ساده و چندشاخگی در ترک با دقت ۹۹/۳ درصد انجام داد [۱۰]. خلاصه‌ای از

در دهه اخیر، تشخیص و طبقه‌بندی ترک‌های سطحی با استفاده از تکنیک‌های پردازش تصویر و یادگیری ماشین رونق گرفته است. در سال ۲۰۱۴ یک روش شناسایی و طبقه‌بندی خودکار ترک برای کنترل ایمنی تونل مترو ارائه شد که بر اساس روش آستانه‌گذاری پیشرفته بود و دقت بیش از ۹۰ درصد را به همراه داشت [۲۷]. در سال ۲۰۱۵ یک سیستم خودکار برای شناسایی و طبقه‌بندی ترک‌های سطحی بر اساس فیلتر مبتنی بر ماتریس هسین طراحی شد که دقت نتایج بدست آمده بیش از ۹۰ درصد بود [۲۸]. یک روش تشخیص و طبقه‌بندی ترک‌های سطحی پل‌ها بر اساس یک ربات کوهنوردی با استفاده از الگوریتم ماشین بردار پشتیبانی در سال ۲۰۱۶ مطرح شد [۲۹]. در روش مذکور از فیلتر وینر، تبدیل موجک و آنالیز مورفولوژی برای پیش‌پردازش استفاده شد و میانگین دقت آن ۹۱/۸۵ درصد بود. مطالعه‌ای دیگر که در سال ۲۰۱۷ انجام شد الگوریتمی برای طبقه‌بندی ترک‌ها به منظور مدیریت بهتر آسفالت‌ها ارائه داد که در آن پس از پردازش تصاویر ترک‌ها، از الگوریتم درخت

ماتریسی) بر روی داده سه بعدی صورت بگیرد. حال اگر داده دو بعدی موجود باشد و معکوس ماتریس قبل در آن ضرب شود، ویژگی‌های سه بعدی آن تصویر در دسترس خواهد بود که این روند همان روند شناسایی اشیاست. استفاده از این ایده در عمل داده‌های لازم برای آموزش شبکه‌ها را بسیار کاهش خواهد داد. این راهکار تحت عنوان شبکه‌های کپسولی شناخته می‌شود.

از طرف دیگر در کاربردهای عملی، مهم است که ترک‌های سطحی موجود در تصاویری که حاوی ترک‌های واقعی هستند از تصاویری که الگوهای نویزی شبیه به ترک سطحی (مانند سوراخ‌ها، توده‌ها، لکه‌ها و سایه‌های تاریک) دارند که اغلب هم در سازه‌های بتنی دیده می‌شوند، تشخیص داده شوند. بعلاوه اگر تشخیص چندشاخگی در ترک‌های سطحی نیز به این موارد اضافه شود، موضوع پیچیده‌تر می‌شود. این ویژگی‌ها باعث می‌شوند که تشخیص ترک‌های سطحی به آسانی تشخیص یک شیء در یک تصویر نباشد و از آنجا که این مجموعه تصاویر در پژوهش‌های قبلی کمتر مورد توجه نویسندگان بوده‌اند، تصاویر ترک‌های سطحی در سازه‌های بتنی می‌تواند یک داده مناسب و مفید جهت ارزیابی استراتژی‌های طبقه‌بندی نامتوازن تصاویر باشند.

الگوریتم‌های استفاده‌شده برای طبقه‌بندی ترک‌ها در جدول ۲ مشاهده می‌گردد.

بررسی مطالعات پیشین در زمینه طبقه‌بندی ترک‌های سطحی نشان می‌دهد که الگوریتم‌های یادگیری عمیق نسبت به سایر الگوریتم‌ها به مراتب کارآمدتر و موثرتر عمل کرده‌اند. به عنوان نمونه شبکه طراحی شده در [۱۰] توانسته است ترک‌های موجود در سازه‌های بتنی را با صحت ۹۹/۳ درصد طبقه‌بندی کند. اما نکته قابل تامل این است که اکثر الگوریتم‌های یادگیری عمیق مانند شبکه‌های عصبی کانولوشنی، به تعداد بسیار زیادی از تصاویر برای مدل‌سازی مناسب در مرحله آموزش نیاز دارند. این ویژگی باعث می‌شود که این الگوریتم‌ها در طبقه‌بندی مجموعه نامتوازن تصاویر (مثلاً تشخیص چندشاخگی در ترک‌های سطحی) عملکرد مناسبی نداشته باشند. هم‌چنین وجود تعداد تصاویر بالا در مدل‌سازی الگوریتم‌های یادگیری عمیق منجر به الزام وجود سیستم‌های سخت‌افزاری قوی می‌شود که همین نکته استفاده از آنها را با محدودیت‌هایی مواجه می‌کند. راه‌حل مقابله با این مسائل، استفاده از معکوس گرافیک کامپیوتری است. ایده گرافیک کامپیوتری به این صورت است که در آن برای رسیدن از یک داده سه بعدی به یک داده دو بعدی (تصویر) کافی است که یک تغییر خطی (ضرب

جدول ۲ خلاصه‌ای از الگوریتم‌های استفاده‌شده در طبقه‌بندی ترک‌های سطحی در سازه‌های بتنی

روش	سال	نام دسته‌ها	صحت (درصد)
آستانه‌گذاری پیشرفته [۲۷]	۲۰۱۴	بدون ترک، ترک	۹۰
ماتریس هسین [۲۸]	۲۰۱۵	ترک کم، ترک متوسط، ترک زیاد	۹۰
ماشین بردار پشتیبانی [۲۹]	۲۰۱۶	طولی، عرضی، بلوکی، شبکه‌ای	۹۱/۸۵
درخت تصمیم‌گیری [۳۰]	۲۰۱۷	بدون ترک، ترک	۸۰
یادگیری عمیق [۳۱]	۲۰۱۸	بدون ترک، ترک	۹۲
شبکه تماماً رمزگذار-رمزگشا [۳۲]	۲۰۱۹	بدون ترک، ترک	۹۲
شبکه‌های عصبی کانولوشنی [۳۳]	۲۰۲۰	بدون ترک، ترک	۹۴
شبکه‌های عصبی کانولوشنی [۱۰]	۲۰۲۰	بدون ترک، ترک ساده، چندشاخگی در ترک	۹۹/۳

ترک، دسته اقلیت را تشکیل می‌دهند. نوآوری‌ها و دستاوردهای این مقاله عبارتند از:

- ۱- جمع‌آوری و معرفی مجموعه تصاویر جدید برای ترک‌های سطحی در سازه‌های بتنی.
- ۲- طراحی یک معماری جدید مبتنی بر شبکه‌های عصبی کپسولی برای طبقه‌بندی تصاویر.
- ۳- ارزیابی عملکرد معماری پیشنهادی در طبقه‌بندی متوازن و نامتوازن تصاویر.
- ۴- طبقه‌بندی متوازن مجموعه تصاویر جمع‌آوری شده با دقت ۹۹/۵۶ درصد که نسبت به شبکه‌های عصبی کانولوشنی و سایر روش‌ها بالاتر است.
- ۵- عملکرد مناسب معماری پیشنهادی در طبقه‌بندی نامتوازن مجموعه تصاویر جمع‌آوری شده تا عدم توازن دسته اقلیت به اکثریت ۱ به ۸ که دارای دقت بالای ۸۰ درصد است.

با توجه به موارد ذکر شده در پاراگراف‌های پیشین، برای شناسایی و طبقه‌بندی مؤثر چندشاخگی در ترک‌های سطحی که در مقایسه با سایر ترک‌های سطحی یک دسته اقلیت به شمار می‌رود و به‌منظور تولید یک مدل طبقه‌بندی قوی نیاز به استخراج ویژگی‌های پیشرفته از تصاویر ترک‌های سطحی در سازه‌های بتنی است. از طرفی ویژگی‌های تصاویر ترک‌های سطحی در سازه‌های بتنی آنها را به یک مجموعه داده مناسب جهت ارزیابی مدل‌های طبقه‌بندی نامتوازن کرده است. به همین منظور در این پژوهش، یک روش هوشمند و ترکیبی بر اساس شبکه‌های عصبی کپسولی برای طبقه‌بندی متوازن و نامتوازن مجموعه تصاویر ترک‌های سطحی در سازه‌های بتنی ارائه می‌شود که در نتیجه آن ترک‌های سطحی به سه دسته بدون ترک، ترک ساده و چندشاخگی در ترک طبقه‌بندی می‌شوند. در طبقه‌بندی نامتوازن این تصاویر، دسته‌های بدون ترک و دارای ترک ساده، دسته‌های اکثریت و دسته چندشاخگی در

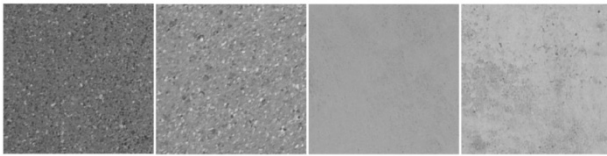
۶- مقاومت مناسب معماری پیشنهادی در کاهش تعداد داده‌ها در مقایسه با شبکه‌های کانولوشنی. به این ترتیب و در ادامه در بخش دوم، معرفی و آماده‌سازی مجموعه داده صورت می‌پذیرد. بخش سوم به مواد و روش‌ها شامل شبکه‌های کانولوشنی، شبکه‌های کیسولی و شبکه کیسولی پیشنهادی اختصاص دارد. در بخش چهارم نتایج در قالب سه بخش و با استفاده از شکل‌ها و جدول‌ها بیان می‌شوند. در نهایت جمع‌بندی در بخش پنجم صورت می‌گیرد.

۲ معرفی و آماده‌سازی مجموعه داده

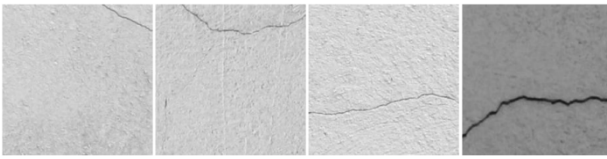
همان‌گونه که پیش‌تر ذکر شد مجموعه تصاویر ترک‌های سطحی در سازه‌های بتنی به دلیل پیچیدگی‌های موجود در ساختارشان و اینکه کمتر مورد توجه مقالات و پژوهش‌های پیشین در بحث طبقه‌بندی مجموعه نامتوازن تصاویر قرار گرفته‌اند، مجموعه تصاویر مناسب و مفیدی در این حوزه به شمار می‌روند. در این پژوهش سازه‌های بتنی اعم از ساختمان‌ها و سوله‌های صنعتی با بتن از جنس سیمان نوع ۲ و عیار ۲۵۰ تا ۳۰۰ برای جمع‌آوری تصاویر در نظر گرفته شدند. ۹۰۰ تصویر تهیه‌شده توسط یک پهپاد به صورت تصاویر با کیفیت 4K ذخیره شدند. شکل ۱ یک نمونه از سازه بتنی مورد استفاده در این پژوهش را نشان می‌دهد. به دلیل اینکه تصاویر به صورت تصاویر رنگی ذخیره می‌شوند، لازم است ابتدا به تصاویر خاکستری تبدیل می‌شوند. سپس، هر کدام از ۹۰۰ تصویر خاکستری حاصل به تصاویر کوچک با ابعاد 330×330 تقسیم می‌شوند. در نهایت ۱۳۵۰۰ تصویر برای ارزیابی مرحله طبقه‌بندی انتخاب شد. شکل‌های ۲، ۳ و ۴ نمونه تصاویر با ابعاد 330×330 را نشان می‌دهند که این شکل‌ها به ترتیب متعلق به دسته تصاویر بدون ترک، تصاویر دارای ترک ساده و تصاویر دارای چندشاخگی در ترک هستند. هم‌چنین به دلیل وجود تغییر رنگ یا محوشدگی در بین برخی از تصاویر ایجاد شده، لازم است پیش‌پردازشی جهت رفع این موارد بر روی این تصاویر صورت گیرد. این عمل با استفاده از تعدیل هیستوگرام انجام می‌شود. شکل ۵ (الف) نمونه‌ای از تصاویر با کیفیت پایین را نشان می‌دهد. شکل ۵ (ب) حاصل تعدیل هیستوگرام روی این تصویر را به نمایش گذاشته است. همان‌طور که قسمت‌های (ج) و (د) این شکل نشان می‌دهند، تعدیل هیستوگرام باعث گسترده شدن هیستوگرام تصویر شده است که خود افزایش کنتراست تصویر را در پی دارد.



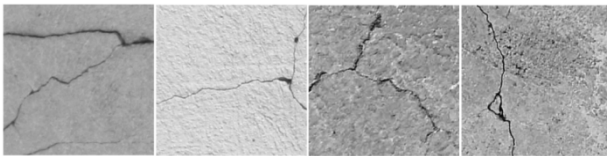
شکل ۱ یک نمونه از سازه بتنی مورد استفاده در این مقاله



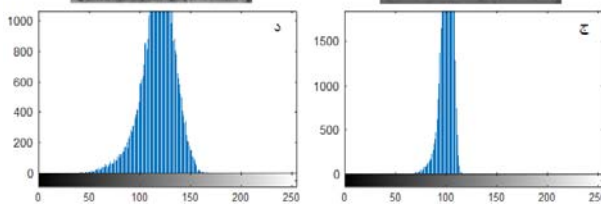
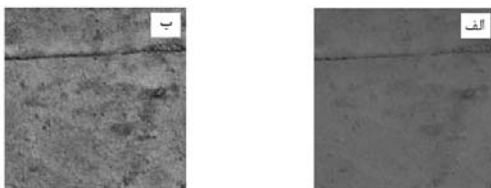
شکل ۲ نمونه‌ای از تصاویر بدون ترک



شکل ۳ نمونه‌ای از تصاویر دارای ترک ساده



شکل ۴ نمونه‌ای از تصاویر دارای چندشاخگی در ترک



شکل ۵ مراحل تعدیل هیستوگرام در یک تصویر نمونه، (الف) تصویر نمونه، (ب) تصویر حاصل از تعدیل هیستوگرام قسمت (الف)، (ج) نمودار هیستوگرام قسمت (الف)، (د) نمودار هیستوگرام قسمت (ب)

۳ مواد و روش‌ها

۳-۱ شبکه‌های عصبی کانولوشنی

شبکه‌های عصبی کانولوشنی^۱ (CNN) نوع خاصی از شبکه‌های عصبی عمیق پیش‌خور هستند که با اعمال تغییراتی بر شبکه‌های دسته یک پرسپترون چندلایه طراحی شده‌اند و انتخابی مناسب برای پردازش تصاویر هستند [۳۴]. در این شبکه‌ها، لایه‌های ابتدایی‌تر وظیفه دارند تا ویژگی‌های ابتدایی را استخراج کنند (مانند گرادیان رنگ‌ها و لبه‌های تصویر) و در لایه‌های بالاتر با ترکیب آن‌ها ویژگی‌های با سطوح بالاتر بسازند [۱۰]. این ترکیب شدن، با جمع وزن‌دار ویژگی‌های پایین‌تر و عبور از یک تابع غیرخطی اتفاق می‌افتد. اما این شبکه‌ها ایراداتی نیز دارند که عبارتند از: عدم در نظر گرفتن ارتباطات فضایی ویژگی‌های استخراج شده، از بین رفتن بخشی از داده‌ها در لایه‌های ادغام و عدم پاسخگویی مناسب این شبکه‌ها به تصاویر ورودی از زاویه دیدهای مختلف. در حالت کلی اگر تصویر ورودی از یک زاویه

¹ Convolutional Neural Network

۲- وزن‌دهی عددی به بردارهای ایجادشده در مرحله قبل، جمع وزن‌دار بردارها و غیرخطی‌سازی برداری.

الف) ضرب ماتریسی بردارهای ورودی با ضرایب وزنی: شکل ۷ ساختارهای مربوط به نورون و کپسول را نشان می‌دهد. بر اساس قسمت (ب) این شکل، بردارهای ورودی که کپسول دریافت می‌کند (u_1, u_2, u_3) از سه کپسول دیگر در لایه و سطح پایین آمده‌اند. طول این بردارها نشان‌دهنده احتمالاتی است که کپسول‌های سطح پایین اشیاء مربوط به خود را شناسایی کرده‌اند و جهت بردارها برخی از حالت‌های داخلی اشیاء کشف شده را نشان می‌دهد. سپس، این بردارها با ماتریس‌های وزنی w ضرب می‌شوند که ارتباطات مکانی و مهم دیگری را بین ویژگی‌های سطح پایین (u_1, u_2, u_3) و ویژگی سطح بالا (u_j) نشان می‌دهند. خروجی حاصل شده، موقعیت پیش بینی شده از ویژگی سطح بالا است که مطابق معادله (۱) ساخته خواهد شد [۳۷]:

$$\hat{u}_{j|i} = w_{ij} u_i \quad (1)$$

ب) وزن‌دهی عددی به بردارهای ایجادشده در مرحله قبل، جمع بردارهای وزن‌دهی شده و غیرخطی‌سازی برداری: پس از مشخص شدن $\hat{u}_{j|i}$ ها، برای یافتن c از الگوریتم مسیریابی پویا استفاده می‌شود که یک روش جدید برای تعیین محل خروجی هر کپسول است [۳۷]. در واقع در این مرحله هدف آن است که تصمیم گرفته شود که کپسول‌های لایه قبل‌تر به کدام کپسول لایه بالاتر متعلق باشد. به بیانی دیگر مقادیر c به این منظور وجود دارند که مشخص کنند هر کپسول لایه پایین چه میزان مرتبط با کپسول‌های لایه بعد هستند. مراحل الگوریتم مسیریابی پویا در جدول ۳ ذکر شده است. مطابق این جدول، ضرایب c_{ij} نشان‌دهنده آن است که کپسول i ام از لایه l با چه احتمالی به کپسول z ام از لایه $l+1$ متعلق است. تعداد تکرارها به وسیله r نمایش داده می‌شود. منظور از l لایه قبل کپسول است. مقدار b مقدار متغیر محلی در معادلات است که برای به روزکردن وزن‌های c به کار برده می‌شود و در ابتدا نیز مقدار صفر به آن اختصاص داده شده است. در مرحله (۲-۱) این جدول از تابع $softmax$ استفاده شده است. معادله (۲)، معادله مربوط به $softmax$ را نشان می‌دهد [۳۷]:

$$\sigma: \mathbb{R}^k \rightarrow \left\{ z \in \mathbb{R}^k \mid z_i \geq 0, \sum_{i=1}^k z_i = 1 \right\}, \quad \sigma_z = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}}, \quad j = 1, \dots, K \quad (2)$$

در مرحله (۲-۲) جدول ۳ از جمع وزن‌دار کپسول‌های لایه قبل، کپسول لایه بعد به دست می‌آید و سپس در مرحله (۳-۲) مقدار آن وارد تابع فعال‌ساز غیرخطی می‌شود. نوآوری دیگری که در شبکه‌های کپسولی معرفی شده، یک تابع فعال‌سازی غیرخطی جدید است که یک بردار را دریافت کرده و سپس آن را لهیده

دیگر گرفته شود و شبکه، آموزش مشابه آن را ندیده باشد، شبکه CNN قادر به شناسایی نخواهد بود و تنها می‌تواند در برابر تغییرات تصویر در یک زاویه دید خاص مقاوم باشد. این ویژگی باعث می‌شود که این شبکه‌ها به داده‌های بسیار زیادی برای آموزش شناسایی نیاز داشته باشند. همان‌گونه که پیش‌تر اشاره شد، راه‌حل مقابله با این مسائل، استفاده از معکوس گرافیک کامپیوتری در قالب شبکه‌های عصبی کپسولی است که در ادامه با جزئیات بیشتر معرفی می‌گردند.

۲-۳ شبکه‌های عصبی کپسولی

شبکه عصبی کپسولی^۱ (CapsNet) یک سیستم یادگیری ماشین است که می‌تواند برای مدل کردن بهتر روابط سلسله مراتبی مورد استفاده قرار گیرد [۳۵]. کار اصلی این شبکه‌ها، اضافه کردن ساختارهایی به نام کپسول به شبکه‌های CNN و استفاده مجدد خروجی آن برای کپسول‌های مرتبه بالاتر برای رسیدن به فرم باثبات‌تر است [۳۶]. شکل ۶ ساختار کلی شبکه‌های CapsNet را نشان می‌دهد که دارای دو بخش است: بخش رمزگذار و بخش رمزگشا. بخش رمزگذار معمولاً از سه لایه کانولوشن، کپسول‌های اولیه و کپسول‌های دسته‌ها تشکیل می‌شود. بخش رمزگذار، تصاویر را به عنوان ورودی می‌گیرد و طی آموزش یاد می‌گیرد که آن‌ها را درون بردارهایی از پارامترها رمزگذاری کند. بخش رمزگشا که از لایه‌های تماماً متصل تشکیل شده است، یک بردار را از کپسول دسته‌ها می‌گیرد و می‌آموزد که آن را به صورت یک تصویر رمزگشایی کند. ذکر این نکته ضروری است که رمزگشا تنها بردار کپسول دسته صحیح را در طول پروسه آموزش در نظر می‌گیرد و از موارد غیر صحیح چشم‌پوشی می‌کند. در واقع بخش رمزگشا به عنوان تنظیم‌کننده مورد استفاده قرار می‌گیرد. به طوری که خروجی صحیح کپسول دسته‌ها را به عنوان ورودی دریافت می‌کند و یاد می‌گیرد که یک تصویر را متناظر با آن بازسازی کند. تابع هزینه این کار، فاصله اقلیدسی بین تصویر بازسازی شده و تصویر ورودی است.

برای درک ساختار کپسول لازم است ابتدا مرور کوتاهی بر ساختارهای قبلی شبکه‌های عمیق شود. خروجی هر نورون در شبکه‌های عصبی یک اسکالر (یک عدد اعشاری) است. مقادیر خروجی در شبکه‌های CapsNet به صورت یک بردار ذخیره می‌شود؛ در حالی که مقادیر در شبکه‌های CNN به صورت عددی است. ویژگی بردار با خود خاصیت‌های مهمی از جمله اندازه و جهت دارد. به طوری که وقتی ویژگی‌ها در هر جای تصویر تغییر می‌کنند، اندازه بردار حفظ می‌شود اما جهت آن تغییر می‌کند. در حالت کلی یک کپسول از دو بخش مهم تشکیل می‌گردد که در ادامه به توضیح این بخش‌ها پرداخته می‌شود [۳۷]:

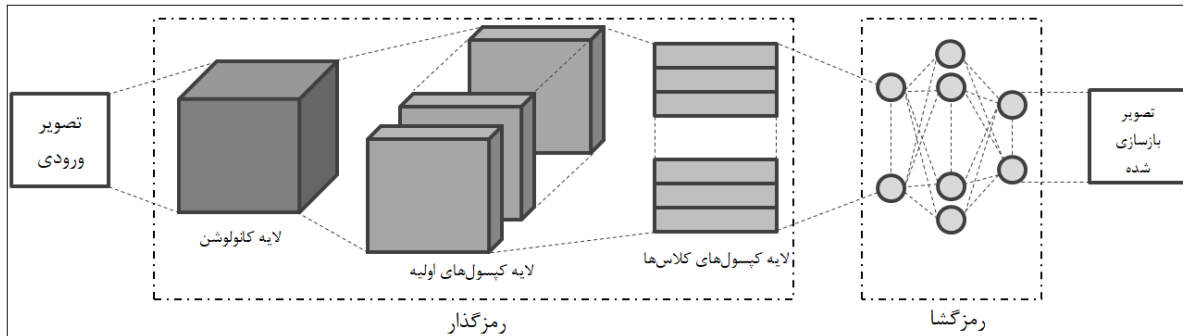
۱- ضرب ماتریسی بردارهای ورودی با ضرایب وزنی.

¹ Capsule Neural Network

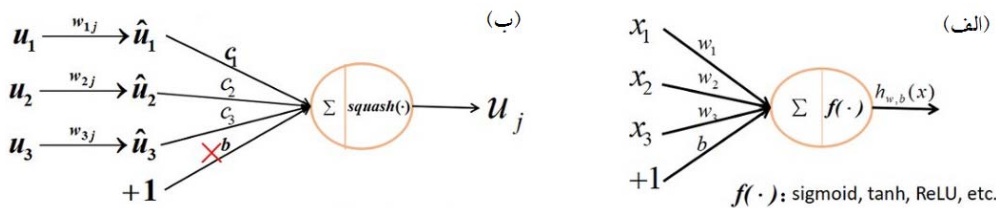
در آخر وزن‌های مربوط به b و c به‌روزرسانی می‌شوند. برای این کار از میزان شباهت استفاده می‌شود، یعنی اگر بردار تولیدشده از مرحله قبل (پس از ضرب وزن‌دار) در راستای کپسول لایه بعد باشد، باید وزن بیشتر به c_{ij} مربوط اختصاص داده شود که بهترین معیار برای بررسی میزان شباهت دو بردار نیز ضرب داخلی است. این مراحل بایستی r بار تکرار شود.

می‌کند که طول آن بیش از یک نباشد، ولی جهت بردار را حفظ می‌کند. این یک روش نوین غیرخطی‌سازی بردار-به-بردار است که توسط این تابع صورت می‌پذیرد [۳۸]:

$$u_j = \frac{\|s_j\|^2}{1 + \|s_j\|^2} \times \frac{s_j}{\|s_j\|} \quad (3)$$



شکل ۶ ساختار کلی شبکه‌های عصبی کپسولی



شکل ۷ ساختارهای مربوط به (الف) نرون، (ب) کپسول

جدول ۳ مراحل الگوریتم مسیریابی پویا [۳۷]

الگوریتم $(\hat{u}_{j i}, r, l)$ Routing		
$b_{ij} \rightarrow \cdot$	برای کپسول i ام در لایه l و کپسول j ام در لایه $(l+1)$	۱
	برای تکرار r :	۲
$\text{softmax}(b_i) \rightarrow c_i$	برای کپسول i ام در لایه l	۱-۲
$\sum_i c_{ij} \hat{u}_{j i} \rightarrow s_j$	برای کپسول j ام در لایه $(l+1)$	۲-۲
$\text{squash}(s_j) \rightarrow u_j$	برای کپسول j ام در لایه $(l+1)$	۳-۲
$b_{ij} + \hat{u}_{j i} u_j \rightarrow b_{ij}$	برای کپسول i ام در لایه l و کپسول j ام در لایه $(l+1)$	۴-۲
	بازگرداندن u_j	۵-۲

طول پیش‌بینی بردارهای ۳ بعدی از طول خروجی‌های کپسول‌های لایه کپسول‌های ترک است.

در جدول ۴ جزئیات لایه‌های بخش رمزگذار و رمزگشا ثبت شده است. بر اساس این جدول، لایه‌های کانولوشن اول تا سوم به ترتیب از ۳۲، ۶۴ و ۹۶ فیلتر با گام ۲ و ابعاد مشخص تشکیل شده است. تابع فعال‌ساز این سه لایه کانولوشنی تابع $ReLU$ است. لایه چهارم که لایه کانولوشن ترکیبی نام دارد، از یک لایه ادغام سراسری، یک لایه کانولوشن با تابع فعال‌ساز $ReLU$ ، یک لایه کانولوشن با تابع فعال‌ساز \tanh و یک قسمت ترکیب‌کننده تشکیل شده است. شکل ۹ جزئیات لایه کانولوشن ترکیبی را نشان می‌دهد. در این لایه ابتدا ادغام سراسری به ۹۶ نگاشت ویژگی ورودی

۳-۳ ساختار CapsNet پیشنهادی

شبکه طراحی شده در این پژوهش در شکل ۸ مشاهده می‌گردد. قسمت رمزگذار از سه لایه کانولوشن، یک لایه کانولوشن ترکیبی، کپسول‌های اولیه و لایه کپسول‌های ترک تشکیل شده است. سه لایه تماماً متصل که از خروجی لایه کپسول‌های ترک تغذیه می‌شوند، قسمت رمزگشا را تشکیل می‌دهند. تصاویر سازه‌های بتنی با ابعاد 330×330 به عنوان ورودی وارد بخش رمزگذار شبکه می‌شوند و این بخش طی آموزش یاد می‌گیرد که آن‌ها را درون بردارهای ۱۶ بعدی از پارامترها رمزگذاری کند. اینجا جایی است که کپسول‌ها کار خود را انجام می‌دهند. خروجی شبکه در

خروجی لایه کپسول‌های ترک، ۳ بردار ۱۶ بعدی است. در طول پروسه آموزش برای هر نمونه آموزشی یک مقدار هزینه برای هر یک از این ۳ بردار طبق معادله (۹) محاسبه می‌شود [۳۷]:

$$L_c = T_c \cdot \max(\cdot, m^+ - \|V_c\|)^2 + \lambda(1 - T_c) \cdot \max(\cdot, \|V_c\| - m^-)^2 \quad (9)$$

که V_c ($c = 1, 2, 3$) هر یک از بردارهای خروجی، λ ، m^+ و m^- پارامترهای ثابت هستند. برای محاسبه هزینه نهایی هر نمونه، سه مقدار محاسبه شده باهم جمع می‌شوند. از آنجا که در اینجا یادگیری نظارت شده وجود دارد، هر نمونه در مجموعه داده آموزشی دارای برچسب صحیح است که در این حالت یک بردار سه بعدی با دو عنصر صفر و یک عنصر با مقدار یک (موقعیت صحیح) وجود دارد. این شیوه به کدبندی One-Hot معروف است [۳۹]. در فرمول تابع هزینه، برچسب صحیح، مقدار T_c را مشخص می‌کند: این مقدار چنانچه برچسب صحیح به کپسول ترک خاص مربوط باشد برابر یک خواهد بود و در غیر آن صورت برابر صفر است. در لایه آخر (لایه کپسول‌های ترک) کپسول‌های اول، دوم و سوم به ترتیب مسئول رمزگذاری تصاویر بدون ترک، تصاویر دارای ترک ساده و تصاویر دارای چندشاخگی در ترک هستند. اگر فرض شود که برچسب صحیح برابر دسته تصاویر بدون ترک باشد، این نشان می‌دهد که کپسول نخست در لایه پنجم مسئول رمزگذاری آن خواهد بود. برای این کپسول ترک، مقدار T_c برابر یک بوده و برای دو کپسول باقی مانده در این لایه، برابر صفر خواهد بود. وقتی T_c برابر یک است، بخش اول تابع هزینه محاسبه می‌شود و بخش دوم برابر صفر خواهد بود. در شبکه پیشنهادی این مقاله، $m^+ = 0.9$ ، $m^- = 0.1$ و $\lambda = 0.5$ است. قسمت رمزگشا نیز مطابق شکل ۸ یک بردار ۱۶ بعدی را از کپسول ترک صحیح در لایه آخر بخش رمزگذار می‌گیرد و می‌آموزد که آن را به صورت یک تصویر ترک رمزگشایی کند. طبق جدول ۴ خروجی این لایه پس از تغییر شکل، تصویر رمزگشایی شده با ابعاد 330×330 را به خروجی تحویل می‌دهد.

اعمال می‌شود. نگاشت‌های ویژگی ورودی این لایه دارای ابعاد 38×38 هستند. در حالت کلی اگر ورودی این لایه V_{c_p} با ابعاد $M \times N \times Q$ باشد (Q نگاشت ویژگی با ابعاد $M \times N$)، ادغام سراسری برای هر نگاشت ویژگی طبق رابطه (۴) محاسبه می‌گردد [۳۵]:

$$V_{gq_k} = \sum_{i=1}^M \sum_{j=1}^N q_{kij} \quad (4)$$

که q_k درایه‌های نگاشت‌های ویژگی است. اگر V_g خروجی لایه ادغام سراسری با ابعاد $1 \times 1 \times Q$ باشد، خروجی‌های لایه‌های کانولوشن با تابع فعال‌ساز $ReLU$ و $tanh$ عبارت است از:

$$V_1 = ReLU(\omega_1 V_g + b_1) \quad (5)$$

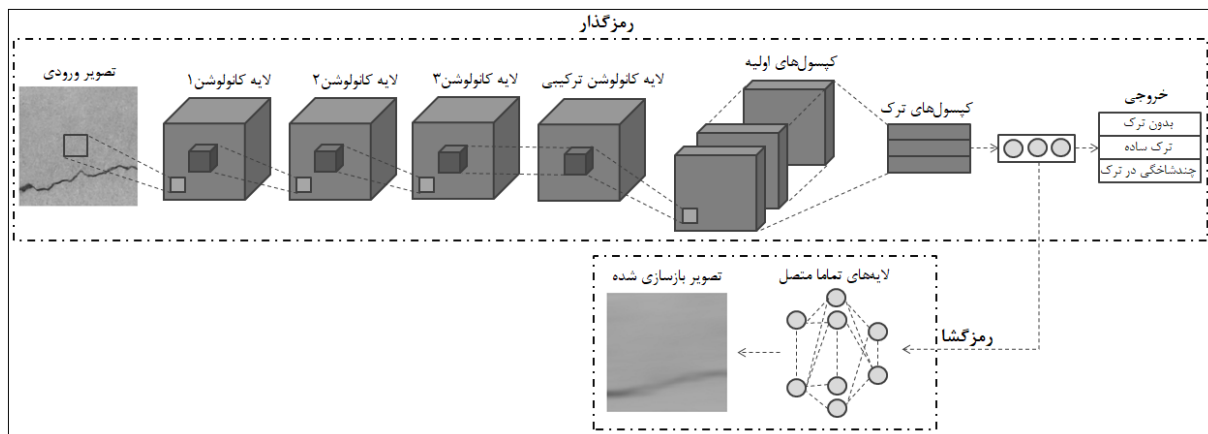
$$V_2 = ReLU(\omega_2 V_1 + b_2) \quad (6)$$

که ω_1 و ω_2 ماتریس‌های وزنی و b_1 و b_2 ضرایب ثابت متناظر هستند. در نهایت خروجی این لایه از رابطه (۸) بدست می‌آید (*): علامت کانولوشن است):

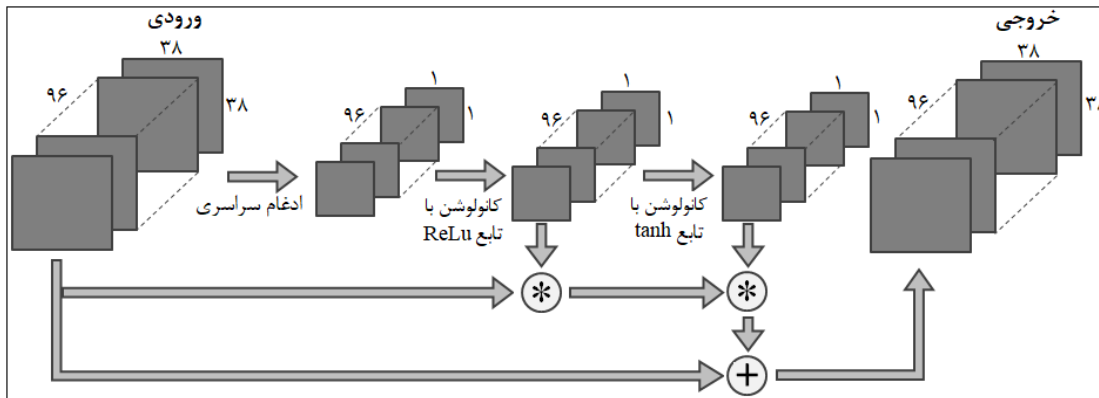
$$V_3 = V_{c_p} * V_1 \quad (7)$$

$$V_{ct} = V_{c_p} + (V_3 * V_2) \quad (8)$$

لایه کپسول‌های اولیه دارای ۳۲ کپسول است که وظیفه آن دریافت ویژگی‌های اساسی است که توسط لایه‌های کانولوشن قبلی کشف شده و ایجاد ترکیبی از این ویژگی‌هاست. هر کپسول در این لایه از ۸ فیلتر با ابعاد 9×9 و گام ۲ تشکیل می‌شود. از آنجا که ورودی این لایه، نگاشت‌های ویژگی با ابعاد $38 \times 38 \times 128$ می‌باشد، نگاشت‌های ویژگی با ابعاد $16 \times 16 \times 8$ خروجی هر کپسول را تشکیل می‌دهند. لایه آخر (لایه کپسول‌های ترک) دارای سه عدد کپسول است که برای هر دسته از تصاویر یک کپسول در نظر گرفته شده است. هر کپسول به عنوان ورودی یک نگاشت ویژگی با ابعاد $16 \times 16 \times 8 \times 32$ را در نظر می‌گیرد. در واقع ورودی هر کپسول در این لایه تعداد $16 \times 16 \times 32 = 8192$ بردار ۸ بعدی است. طبق اساس کار داخلی کپسول، هر یک از این بردارهای ورودی، ماتریس وزنی خود با ابعاد 8×16 را می‌گیرد تا فضای ۸ بعدی ورودی کپسول به فضای ۱۶ بعدی خروجی کپسول نگاشت شود.



شکل ۸ ساختار شبکه CapsNet پیشنهادی



شکل ۹ جزئیات لایه کانولوشن ترکیبی در شبکه CapsNet پیشنهادی

جدول ۴ جزئیات لایه‌های قسمت رمزگذار و رمزگشا

بخش	نام لایه	ورودی لایه	تعداد (فیلتر/کپسول/نورون)	خروجی لایه
رمزگذار	کانولوشن ۱	تصویر ۳۳۰×۳۳۰	۳۲ فیلتر ۵×۵ با گام ۲	نگاشت‌های ویژگی ۱۶۴×۱۶۴×۳۲
	کانولوشن ۲	نگاشت‌های ویژگی ۱۶۴×۱۶۴×۳۲	۶۴ فیلتر ۷×۷ با گام ۲	نگاشت‌های ویژگی ۸۰×۸۰×۶۴
	کانولوشن ۳	نگاشت‌های ویژگی ۸۰×۸۰×۶۴	۹۶ فیلتر ۷×۷ با گام ۲	نگاشت‌های ویژگی ۳۸×۳۸×۹۶
	کانولوشن ترکیبی	نگاشت‌های ویژگی ۳۸×۳۸×۹۶	ادغام سراسری ۹۶ فیلتر ۱×۱ با گام ۱ و تابع فعال‌ساز ReLu ۹۶ فیلتر ۱×۱ با گام ۱ و تابع فعال‌ساز tanh	نگاشت‌های ویژگی ۳۸×۳۸×۹۶
	کپسول‌های اولیه	نگاشت‌های ویژگی ۳۸×۳۸×۹۶	۳۲ کپسول دارای ۸ فیلتر ۹×۹ با گام ۲	ماتریس‌های ۱۶×۱۶×۸×۳۲
	کپسول‌های ترک	ماتریس‌های ۱۶×۱۶×۸×۳۲	۳ کپسول ۱۶ بعدی	ماتریس‌های ۳×۱۶
	تماما متصل ۱	ماتریس ۳×۱۶	۵۱۲ نورون	۵۱۲ نورون
رمزگشا	تماما متصل ۲	۵۱۲ نورون	۱۰۲۴ نورون	۱۰۲۴ نورون
	تماما متصل ۳	۱۰۲۴ نورون	۱۰۸۹۰۰ نورون	۱۰۸۹۰۰ (۳۳۰×۳۳۰) نورون

۴ نتایج

در این قسمت با توجه شبکه CapsNet پیشنهادی، ابتدا طبقه‌بندی متوازن در ترک‌های سطحی انجام می‌گردد. در ادامه اثرات کاهش کلی تعداد تصاویر بررسی می‌شوند. در نهایت طبقه‌بندی نامتوازن با توجه به دسته تصاویر چندشاخگی در ترک‌های سطحی صورت می‌پذیرد. قابل ذکر است که در هر مرحله مقایسه با سایر الگوریتم‌ها به خصوص شبکه‌های CNN انجام می‌شود.

۴-۱ بررسی طبقه‌بندی متوازن در ترک‌های سطحی

همان‌گونه که پیش‌تر ذکر شد ۱۳۵۰۰ تصویر ترک‌های سطحی از سازه‌های بتنی با ابعاد ۳۳۰×۳۳۰ برای مرحله ارزیابی ایجاد شدند. ابتدا و به منظور مقایسه با محبوب‌ترین شبکه یادگیری عمیق (شبکه CNN)، تعداد ۱۲۰۰۰ تصویر شامل سه دسته تصاویر بدون ترک، تصاویر دارای ترک و تصاویر دارای چندشاخگی در ترک (هر دسته ۴۰۰۰ تصویر) به صورت تصادفی انتخاب شدند. از این تعداد تصویر، تعداد ۹۰۰۰ تصویر (۷۵ درصد مجموعه داده) برای آموزش و ۳۰۰۰ تصویر (۲۵ درصد مجموعه داده) برای آزمایش نیز به صورت تصادفی

گزینه‌ش شدند. ابتدا تصاویر انتخاب شده با استفاده از شبکه CNN طراحی شده در [۱۰] و شبکه CapsNet پیشنهادی طبقه‌بندی و مقایسه شدند. لازم به ذکر است که این طبقه‌بندی‌ها با در نظر گرفتن دسته‌های متوازن (هر دسته ۴۰۰۰ تصویر) انجام شدند. همچنین برای شبکه CapsNet پیشنهادی حالت بدون لایه کانولوشن ترکیبی نیز در نظر گرفته شد. شکل ۱۰ تغییرات درصد صحت در طی مراحل آموزش را با ۹۰۰ تکرار را نشان می‌دهد. طبقه‌بندی تصاویر ترک‌های سطحی توسط شبکه CNN طراحی شده در [۱۰]، شبکه CapsNet پیشنهادی بدون لایه کانولوشن ترکیبی و شبکه CapsNet پیشنهادی به ترتیب میانگین صحت ۹۹/۳، ۹۹/۳۸ و ۹۹/۵۶ درصد را در پی داشت. مطابق انتظار، شبکه‌های یادگیری عمیق با وجود داده‌های کافی و مناسب می‌توانند دقت بالایی در طبقه‌بندی داشته باشند. با این حال در این بررسی، شبکه CapsNet پیشنهادی دقت بالاتری نسبت به دو شبکه دیگر داشته است. همچنین جدول ۵ مقایسه‌ای بین روش‌های طبقه‌بندی ترک‌های سطحی در سازه‌های بتنی با دسته‌های متوازن را نشان می‌دهد. جدول ۶ نیز به بررسی زمانی (در یک سیستم سخت‌افزاری یکسان) بین الگوریتم‌های شناخته‌شده از جمله Alex Net

توسط شبکه CapsNet پیشنهادی نشان می‌دهد. پس از طبقه‌بندی تصاویر ترک به سه دسته در قالب بخش رمزگذار، میتوان در قالب بخش رمزگشا، تصاویر بازسازی شده ورودی را بدست آورد. جدول ۸ نمونه‌ای از تصاویر بازسازی شده به همراه تصاویر اصلی را نشان می‌دهد.

[۴۰]، VGG Net [۴۱]، GoogLe Net [۴۲] و شبکه CapsNet پیشنهادی می‌پردازد که با توجه به آن مشاهده می‌شود که شبکه پیشنهادی توانسته است با صرف زمان کمتر منجر به صحت بالاتری در طبقه‌بندی متوازن ترک‌های سطحی شود. همچنین جدول ۷ نمونه تصاویری را با پیش‌بینی درست و غلط



شکل ۱۰ مقایسه درصد صحت در طبقه‌بندی با دسته‌های متوازن تصاویر ترک‌های سطحی توسط سه شبکه

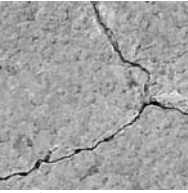
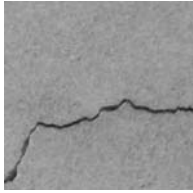
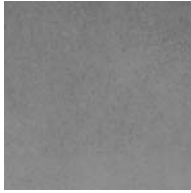

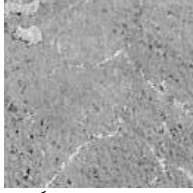
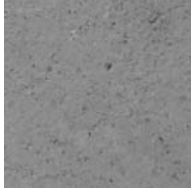
جدول ۵ مقایسه روش‌های طبقه‌بندی ترک‌های سطحی با دسته‌های متوازن

روش	سال	میانگین صحت (درصد)
آستانه‌گذاری پیشرفته [۲۷]	۲۰۱۴	۹۰
ماتریس هسین [۲۸]	۲۰۱۵	۹۰
ماشین بردار پشتیبانی [۲۹]	۲۰۱۶	۹۱/۸۵
درخت تصمیم‌گیری [۳۰]	۲۰۱۷	۸۰
یادگیری عمیق [۳۱]	۲۰۱۸	۹۲
شبکه تماما رمزگذار-رمزگشا [۳۲]	۲۰۱۹	۹۲
شبکه‌های عصبی کانولوشنی [۳۳]	۲۰۲۰	۹۴
شبکه‌های عصبی کانولوشنی [۱۰]	۲۰۲۰	۹۹/۳۱
شبکه عصبی کپسولی پیشنهادی	۲۰۲۰	۹۹/۵۶


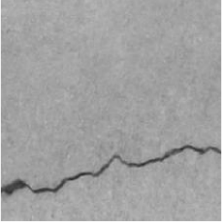
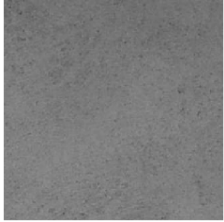



جدول ۶ بررسی زمانی بین برخی از الگوریتم‌های شناخته شده برای طبقه‌بندی ترک‌های سطحی با دسته‌های متوازن

الگوریتم	صحت (درصد)	زمان تقریبی (دقیقه)
Alex Net [۴۰]	۸۱/۳	۱۶۸۹
VGG Net [۴۱]	۸۶/۷	۱۴۲۵
GoogLe Net [۴۲]	۹۴/۱	۴۳۸
شبکه CNN [۱۰]	۹۹/۳۱	۱۲۹
شبکه CapsNet پیشنهادی	۹۹/۵۶	۱۱۷

جدول ۷ نمونه‌ای از تصاویر با پیش‌بینی درست و غلط توسط شبکه CapsNet پیشنهادی

چندشاخگی در ترک	ترک ساده	بدون ترک	دسته واقعی / نوع پیش‌بینی
 (پیش‌بینی شده: چندشاخگی در ترک)	 (پیش‌بینی شده: ترک ساده)	 (پیش‌بینی شده: بدون ترک)	درست
 (پیش‌بینی شده: بدون ترک)	 (پیش‌بینی شده: چندشاخگی در ترک)	 (پیش‌بینی شده: ترک ساده)	غلط

جدول ۸ نمونه‌ای از تصاویر بازسازی شده توسط شبکه CapsNet پیشنهادی

چندشاخگی در ترک	ترک ساده	بدون ترک	دسته / نوع تصویر
			تصویر اصلی
			تصویر بازسازی شده

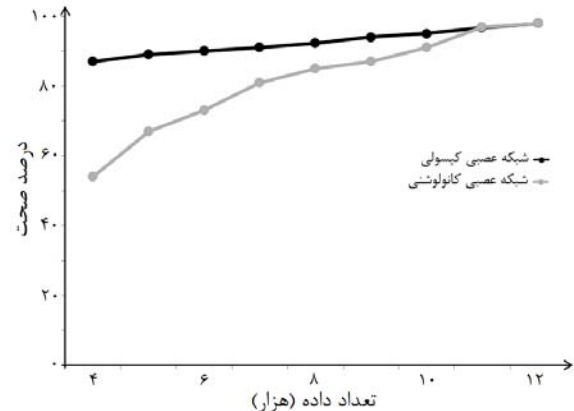
کمتری برای آموزش نیاز باشد. به همین دلیل یک آنالیز حساسیت نسبت به تعداد داده‌ها در شبکه CapsNet پیشنهادی و شبکه CNN طراحی شده در [۱۰] انجام شد. به این صورت که در ۸ مرحله و با کاهش هزارتایی تعداد داده‌ها در هر مرحله اثرات کاهش تعداد داده روی هر دو شبکه بررسی و نتایج به صورت نمودار موجود در شکل ۱۱ به تصویر کشیده شد. همان‌طور که از این شکل دریافت می‌شود با کاهش تعداد داده‌ها گرچه درصد صحت در شبکه کپسولی پیشنهادی شیب نزولی دارد، اما تا محدوده ۶۰۰۰ داده، الگوریتم درصد صحت بالای ۹۰ درصد را حفظ می‌کند که این مورد برای شبکه عصبی کانولوشنی فقط در محدوده ۱۰۰۰۰ داده برقرار است. به تعبیر دیگر، حداقل تعداد داده برای شبکه‌های CapsNet 6000 و برای شبکه‌های CNN 10000 در تحلیل تصاویر ترک‌های سطحی سازه‌های بتنی است. این موضوع تاییدی بر ساختار منحصر به فرد شبکه‌های CapsNet

۲-۴ بررسی اثر کاهش تعداد تصاویر در طبقه‌بندی متوازن در ترک‌های سطحی

در حالت کلی شبکه‌های یادگیری عمیق برای آموزش نیاز به حجم بالایی از داده‌های برجسب خورده دارند. به عبارت دیگر در یادگیری عمیق کافی است تعداد زیادی داده و زمان کافی در اختیار الگوریتم یادگیری قرار داد تا الگوریتم بتواند بهترین ویژگی‌های مناسب به منظور طبقه‌بندی با حداقل خطا را یاد بگیرد. البته همیشه نمیتوان داده‌های برجسب خورده زیادی فراهم کرد و در بسیاری از موارد این موضوع با محدودیت‌هایی (افزایش زمان محاسبات شده و محدودیت‌های سخت‌افزاری) همراه است. از طرف دیگر شبکه‌های CapsNet با در نظر گرفتن بردار ویژگی (به جای مقدار ویژگی در مقایسه با سایر شبکه‌های عصبی عمیق مانند شبکه‌های کانولوشنی) در عمل باعث می‌شوند به تعداد داده‌های

۱). در مرحله اول نسبت دسته اقلیت (تصاویر دارای چندشاخگی در ترک) به دسته اکثریت (تصاویر بدون ترک و دارای ترک ساده) ۱ به ۵ فرض می‌شود. بدین صورت که در هر دسته اکثریت ۴۰۰۰ تصویر و در دسته اقلیت ۸۰۰ تصویر وجود خواهد داشت. این شیوه تا نسبت ۱ به ۱۰ بررسی و درصد صحت دسته اقلیت و کل دسته‌ها ثبت شدند. با توجه جدول ۹، شبکه CapsNet پیشنهادی تا نسبت ۱ به ۸ رفتار مناسبی از خود نشان می‌دهد. در صورتی که شبکه کانولوشنی پیشنهادی در اولین مرحله رفتار مطلوبی ندارد و به عدم توازن دسته‌ها واکنش نشان داده است. نکته دیگری که از این جدول دریافت می‌شود این است که با وجود درصد صحت کلی قابل قبول در شبکه CapsNet در نسبت ۱ به ۱۰ (۸۵/۴۸ درصد)، درصد صحت دسته مهم (دسته اقلیت) مقدار مطلوبی ندارد (۵۹ درصد). این همان موضوع مهمی است که در تحلیل داده‌های نامتوازن به آن پرداخته می‌شود که مناسب بودن درصد صحت کلی در یک شبکه، تضمین‌کننده دقت بالای شبکه در تشخیص دسته و دسته‌های حساس و مهم (دسته اقلیت) نیست. به منظور بررسی کاملتر عملکرد شبکه پیشنهادی، سه مجموعه داده تصویری شناخته شده و معتبر MNIST [۴۳]، CIFAR-10 [۴۴] و Fashion-MNIST [۴۵] انتخاب شدند و نتایج در قالب جدول ۱۰ ثبت شد. لازم به ذکر است که در هر مجموعه داده به طور تصادفی یکی از دسته‌ها به عنوان دسته اقلیت انتخاب شدند. همان‌طور که از جدول ۱۰ مشاهده می‌شود، شبکه CapsNet پیشنهادی در طبقه‌بندی نامتوازن با نسبت دسته اقلیت به اکثریت ۱ به ۱۰ روی چهار مجموعه داده تصویری مختلف توانسته است درصد صحت (مربوط به کل دسته‌ها) بالای ۸۰ درصد را ثبت کند که نشان از عملکرد مناسب شبکه پیشنهادی دارد.

دارد که با استفاده از تبادل دو پارامتر "اندازه" و "جهت" در ویژگی‌های سطح بالا به کاهش تعداد داده‌ها نسبت به شبکه‌های CNN که نوع قدرتمند شبکه‌های عمیق می‌باشد، مقاوم‌تر است.



شکل ۱۱ نمودار تغییرات درصد صحت در شبکه عصبی کیسولی و کانولوشنی بر اساس کاهش تعداد داده‌ها

۳-۴ بررسی طبقه‌بندی نامتوازن در ترک‌های سطحی

در این مرحله با توجه به اینکه شناسایی تصاویر دارای چندشاخگی در ترک از اهمیت بیشتری نسبت به دو دسته دیگر برخوردار است و معمولاً در سازه‌های بتنی یک ناهنجاری و اتفاق نادر به شمار می‌رود، با نامتوازن کردن تصادفی این دسته کارایی شبکه‌های CNN و CapsNet پیشنهادی بررسی می‌شود. برای این منظور مطابق جدول ۹ با ثابت نگه داشتن تعداد داده در دسته تصاویر بدون ترک و دارای ترک ساده و همچنین تغییر تعداد داده در دسته تصاویر چندشاخگی در ترک، درصد صحت هر دسته و درصد صحت کلی بررسی می‌گردد. طبق این جدول، ابتدا ۱۲۰۰۰ تصویر موجود در سه دسته به صورت متوازن وجود دارند (نسبت ۱ به

جدول ۹ بررسی تغییرات نسبت دسته اقلیت به اکثریت در شبکه‌های CapsNet پیشنهادی و CNN

درصد صحت				نسبت دسته اقلیت به اکثریت
شبکه CNN		شبکه CapsNet پیشنهادی		
کل دسته‌ها	دسته اقلیت	کل دسته‌ها	دسته اقلیت	
۹۷/۸۳	۹۸/۱۱	۹۸/۷۱	۹۸/۷	۱ به ۱
۸۴/۷۳	۵۸/۸	۹۷/۳۱	۹۴/۵	۱ به ۵
۷۷/۳۷	۳۶/۷۳	۹۵/۹۳	۹۰/۳۶	۱ به ۶
۷۴/۹۷	۲۹/۵۲	۹۴/۴۸	۸۶/۰۱	۱ به ۷
۷۲/۵۲	۲۲/۱۹	۹۲/۴۸	۸۰	۱ به ۸
۷۱/۴۷	۱۹/۰۳	۸۸/۶۳	۶۸/۴۷	۱ به ۹
۷۱/۱۵	۱۸/۰۸	۸۵/۴۸	۵۹	۱ به ۱۰

جدول ۱۰ نتایج اعمال شبکه CapsNet پیشنهادی روی برخی از مجموعه داده‌ها

مجموعه داده	تعداد کل داده‌ها	تعداد دسته‌ها	نسبت دسته اقلیت به اکثریت	درصد صحت
MNIST [۴۳]	۷۰k	۱۰	۱ به ۱۰	۸۸/۰۱
CIFAR-10 [۴۴]	۶۰K	۱۰	۱ به ۱۰	۸۰/۷۳
Fashion-MNIST [۴۵]	۷۰k	۱۰	۱ به ۱۰	۸۳/۹۶
ترک‌های سطحی	۱۲K	۳	۱ به ۱۰	۸۵/۴۸

بهبودسازی، افزایش داده در دسته اقلیت، کاهش داده در دسته‌های اکثریت و یا ترکیبی از این موارد به مفاهیم شبکه‌های CapsNet در مسائل طبقه‌بندی نامتوازن تصاویر مطالعات بیشتری انجام داد.

۵ جمع‌بندی

چندشاخگی در ترک‌های سطحی یکی از ناهنجاری‌های موجود در سازه‌های بتنی است که شناخت این ناهنجاری می‌تواند به نگهداری سازه‌های بتنی و مدیریت هزینه‌ها کمک شایانی کند. چندشاخگی در ترک‌های سطحی نسبت به ترک‌های ساده، کمتر رخ می‌دهد و این موضوع می‌تواند مساله طبقه‌بندی ترک‌های سطحی را به مسئله طبقه‌بندی نامتوازن تصاویر تبدیل کند. موضوعی که در این مقاله در قالب یک شبکه CapsNet پیشنهادی بررسی شد که نتیجه آن صحت ۹۹/۵۶ درصدی طبقه‌بندی ترک‌های سطحی به سه دسته تصاویر بدون ترک، تصاویر دارای ترک ساده و تصاویر دارای چندشاخگی در ترک بود.

لازم به ذکر است که در طراحی شبکه‌های یادگیری عمیق، مرحله جمع‌آوری داده یکی از مراحل مهم و تقریباً دشوار به شمار می‌رود. چرا که میتوان اساس بازدهی مطلوب یک شبکه عمیق (با معماری مناسب) را تعداد داده‌های آموزش آن شبکه دانست. از این رو ارائه راهکارهایی برای نتیجه‌گیری مناسب با عدم وجود تعداد داده‌های زیاد می‌تواند نقش بسزایی در عملی ساختن الگوریتم‌های تئوری داشته باشد. ساختارهای کپسول موجود در شبکه‌های CapsNet با در نظر گرفتن ارتباطات فضایی ویژگی‌ها، این امکان را فراهم می‌سازند تا با داده‌های به مراتب کمتری نسبت به سایر شبکه‌های عمیق نظیر شبکه‌های CNN، بتوان آموزش شبکه را انجام داد. این موضوع باعث می‌شود که شبکه‌های CapsNet در تعداد داده‌های بالا و یکسان با شبکه‌های CNN نیز نتایج بهتری داشته باشند. همان‌طور که در این مقاله مشاهده شد شبکه CapsNet پیشنهادی نسبت به کاهش داده‌ها بسیار مقاوم‌تر از شبکه‌های CNN است.

بررسی زمانی در یک بستر سخت‌افزاری یکسان نیز نشان از صرف زمان کمتر و درصد صحت بالاتر الگوریتم پیشنهادی در مقایسه با سایر الگوریتم‌های شناخته شده داشت. هم‌چنین بررسی عدم توازن بین دسته‌های تصاویر موجود در ترک‌های سطحی نشان داد که شبکه CapsNet پیشنهادی تا عدم توازن دسته اقلیت به اکثریت ۱ به ۸ رفتار قابل قبولی از خود نشان می‌دهد. حال آنکه این موضوع در شبکه‌های CNN با واکنش منفی روبه‌روست که از نقاط ضعف این شبکه‌ها نیز به شمار می‌رود. هم‌چنین شبکه CapsNet پیشنهادی در سه مجموعه داده مختلف دیگر هم صحت بالای ۸۰ درصد را در حالت طبقه‌بندی نامتوازن حفظ کرد که حاکی از قدرت مناسب این شبکه دارد.

نکته‌ای که در بررسی کاهش داده‌ها و عدم توازن بین دسته‌ها وجود دارد این است که عمل کاهش داده و نامتوازن‌سازی دسته‌ها به صورت تصادفی صورت گرفت. در صورتی که با توجه به ساختارهای شبکه‌های کپسولی، میتوان با حذف هدفمند تصاویر، مقاومت این شبکه‌ها را نسبت به کاهش یا نامتوازی داده‌ها بالا برد. هم‌چنین می‌توان در آینده با تغییراتی نظیر اضافه کردن توابع

مراجع

- [1] S. Hussein, P. Kandel, C. W. Bolan, M. B. Wallace, and U. Bagci, "Lung and pancreatic tumor characterization in the deep learning era: novel supervised and unsupervised learning approaches," *IEEE transactions on medical imaging*, vol. 38, no. 8, pp. 1777–1787, 2019.
- [2] W. Rawat and Z. Wang, "Deep convolutional neural networks for image classification: A comprehensive review," *Neural computation*, vol. 29, no. 9, pp. 2352–2449, 2017.
- [3] I. Cordon, S. Garcia, A. Fernandez, and F. Herrera, "Imbalance: oversampling algorithms for imbalanced classification in R," *Knowledge-Based Systems*, vol. 161, pp. 329–341, 2018.
- [4] جباری، حامد. بیگدلی، نوشین و خادم، علی. "یک روش هیبریدی جدید جهت قطعه‌بندی و تشخیص تومورهای موجود در تصاویر ماموگرافی بافت پستان"، فصلنامه علمی-پژوهشی بیماری‌های پستان ایران، سال نهم، شماره سوم، صفحات ۲۴-۱۴، ۱۳۹۵.
- [5] بیگدلی، نوشین. جباری، حامد و ملکی، نگار. "یک روش ترکیبی هوشمند برای تشخیص، مرزبندی و طبقه‌بندی توده‌های پستان مبتنی بر ویژگی‌های بافت جدید مستخرج از دو نمای تصاویر ماموگرافی"، مجله ماشین‌بینایی و پردازش تصویر سال پنجم، شماره دوم، صفحات ۸۳-۶۹، ۱۳۹۶.
- [6] A. T. Azar, A. Koubaa, N. A. Kamal, and K. M. Fouad, "Improved Dominance Soft Set Based Decision Rules with Pruning for Leukemia Image Classification," *Electronics*, vol. 9, no. 5, p. 794, 2020.
- [7] K. Muhammad, S. Khan, J. Del Ser, and V. H. C. de Albuquerque, "Deep learning for multigrade brain tumor classification in smart healthcare systems: A prospective survey," *IEEE Transactions on Neural Networks and Learning Systems*, 2020.
- [8] Y. He et al., "Deep learning-based classification of primary bone tumors on radiographs: A preliminary study," *EBioMedicine*, vol. 62, p. 103121, 2020.
- [9] N. Vahdati, X. Wang, O. Shirayev, P. Rostron, and F. F. Yap, "External Corrosion Detection of Oil Pipelines Using Fiber Optics," *Sensors*, vol. 20, no. 3, p. 684, 2020.
- [10] بیگدلی، نوشین. جباری، حامد و شجاعی، مهدی. "یک روش هوشمند برای طبقه‌بندی ترک در سازه‌های بتنی بر اساس شبکه‌های عصبی عمیق"، نشریه مهندسی عمران امیرکبیر، ۱۳۹۹.
- [11] V. López, A. Fernández, S. García, V. Palade, and F. Herrera, "An insight into classification with imbalanced data: Empirical results and current trends on using data intrinsic characteristics," *Information sciences*, vol. 250, pp. 113–141, 2013.

- [25] D. Dias-da-Costa, J. Valença, E. Júlio, H. Araújo, and H. Monitoring, "Crack propagation monitoring using an image deformation approach," *Structural Control*, vol. 24, no. 10, p. e1973, 2017.
- [26] A. Mohan and S. Poobal, "Crack detection using image processing: A critical review and analysis," *Alexandria Engineering Journal*, vol. 57, no. 2, pp. 787-798, 2018.
- [27] W. Zhang, Z. Zhang, D. Qi, and Y. Liu, "Automatic crack detection and classification method for subway tunnel safety monitoring," *Sensors*, vol. 14, no. 10, pp. 19307-19328, 2014.
- [28] S. Ghanta, S. S. Shamsabadi, J. Dy, M. Wang, and R. Birken, "A Hessian-based methodology for automatic surface crack detection and classification from pavement images," in *Structural Health Monitoring and Inspection of Advanced Materials, Aerospace, and Civil Infrastructure 2015*, 2015, vol. 9437: International Society for Optics and Photonics, p. 94371Z.
- [29] Y. Chen, T. Mei, X. Wang, and F. Li, "A bridge crack image detection and classification method based on climbing robot," in *2016 35th Chinese Control Conference (CCC)*, 2016: IEEE, pp. 4037-4042.
- [30] A. Cubero-Fernandez, F. J. Rodriguez-Lozano, R. Villatoro, J. Olivares, J. M. Palomares, and V. Processing, "Efficient pavement crack detection and classification," *EURASIP Journal on Image*, vol. 2017, no. 1, pp. 1-11, 2017.
- [31] H. Nhat-Duc, Q.-L. Nguyen, and V.-D. Tran, "Automatic recognition of asphalt pavement cracks using metaheuristic optimized edge detection algorithms and convolution neural network," *Automation in Construction*, vol. 94, pp. 203-213, 2018.
- [32] M. Islam and J.-M. Kim, "Vision-based autonomous crack detection of concrete structures using a fully convolutional encoder-decoder network," *Sensors*, vol. 19, no. 19, p. 4251, 2019.
- [33] B. Li, K. C. Wang, A. Zhang, E. Yang, and G. Wang, "Automatic classification of pavement crack using deep convolutional neural network," *International Journal of Pavement Engineering*, vol. 21, no. 4, pp. 457-463, 2020.
- [34] A. A. M. Al-Saffar, H. Tao, and M. A. Talab, "Review of deep convolution neural network in image classification," in *2017 International Conference on Radar, Antenna, Microwave, Electronics, and Telecommunications (ICRAMET)*, 2017: IEEE, pp. 26-31.
- [35] T. Iesmantas and R. Alzbutas, "Convolutional capsule network for classification of breast cancer histology images," *Image Analysis and Recognition*, pp. 853-860, 2018.
- [36] X. Zhong et al., "An emotion classification algorithm based on SPT-CapsNet," *Neural Computing*, vol. 32, no. 7, pp. 1823-1837, 2020.
- [12] O. Loyola-González, J. F. Martínez-Trinidad, J. A. Carrasco-Ochoa, and M. García-Borroto, "Study of the impact of resampling methods for contrast pattern based classifiers in imbalanced databases," *Neurocomputing*, vol. 175, pp. 935-947, 2016.
- [13] C. Beyan and R. Fisher, "Classifying imbalanced data sets using similarity based hierarchical decomposition," *Pattern Recognition*, vol. 48, no. 5, pp. 1653-1672, 2015.
- [14] P. Branco, L. Torgo, and R. P. Ribeiro, "A survey of predictive modeling on imbalanced domains," *ACM Computing Surveys*, vol. 49, no. 2, pp. 1-50, 2016.
- [15] M. Vannucci and V. Colla, "Genetic algorithms based resampling for the classification of unbalanced datasets," *Intelligent Decision Technologies*, pp. 23-32, 2017.
- [16] L. Yijing, G. Haixiang, L. Xiao, L. Yanan, and L. Jinling, "Adapted ensemble classification algorithm based on multiple classifier system and feature selection for classifying multi-class imbalanced data," *Knowledge-Based Systems*, vol. 94, pp. 88-104, 2016.
- [17] A. Taherkhani, G. Cosma, and T. M. McGinnity, "AdaBoost-CNN: An adaptive boosting algorithm for convolutional neural networks to classify multi-class imbalanced datasets using transfer learning," *Neurocomputing*, vol. 404, pp. 351-366, 2020.
- [18] K. Polat, "A hybrid approach to Parkinson disease classification using speech signal: the combination of smote and random forests," *Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT)*, pp. 1-3, 2019.
- [19] J. Gaspers, Q. Do, and F. Triefenbach, "Data balancing for boosting performance of low-frequency classes in Spoken Language Understanding," *arXiv preprint*, 2020.
- [20] W. M. Hartmann, "Dimension reduction vs. variable selection," *Applied Parallel Computing*, pp. 931-938, 2004.
- [21] V. López, A. Fernández, J. G. Moreno-Torres, and F. Herrera, "Analysis of preprocessing vs. cost-sensitive learning for imbalanced classification. Open problems on intrinsic data characteristics," *Expert Systems with Applications*, vol. 39, no. 7, pp. 6585-6608, 2012.
- [22] A. Ghazikhani, R. Monsefi, H. S. Yazdi, and applications, "Online cost-sensitive neural network classifiers for non-stationary and imbalanced data streams," *Neural computing*, vol. 23, no. 5, pp. 1283-1295, 2013.
- [23] A. Çayır, U. Ünal, H. Dağ, and Security, "Random CapsNet Forest Model for Imbalanced Malware Type Classification Task," *Computers*, p. 102133, 2020.
- [24] W. Zhang, X. Li, X.-D. Jia, H. Ma, Z. Luo, and X. Li, "Machinery fault diagnosis with imbalanced data using deep generative adversarial networks," *Measurement*, vol. 152, p. 107377, 2020.



حامد جباری مدرک کارشناسی ارشد خود را در رشته مهندسی برق گرایش کنترل از دانشگاه بین‌المللی امام خمینی (ره) قزوین در سال ۱۳۹۴ دریافت کرد. ایشان هم‌اکنون دانشجوی دکتری رشته مهندسی برق گرایش کنترل در دانشگاه بین‌المللی امام خمینی (ره) قزوین است. زمینه‌های پژوهشی مورد علاقه وی داده‌کاوی، پردازش تصاویر و یادگیری عمیق است.



نوشین بیگدلی کارشناسی و کارشناسی ارشد مهندسی برق خود را در دانشگاه شیراز به ترتیب در گرایش‌های الکترونیک و کنترل به پایان رساند و در سال ۱۳۸۶ در مقطع دکتری برق-کنترل از دانشگاه صنعتی شریف فارغ‌التحصیل گردید. ایشان هم‌اکنون، دانشیار گروه مهندسی برق-کنترل دانشگاه بین‌المللی امام خمینی (ره) قزوین است. زمینه‌های

تخصصی وی عبارتست از: داده‌کاوی، استخراج ویژگی‌ها و آنالیز سری زمانی، پردازش تصاویر و علائم، سیستم‌های هوشمند، کنترل سیستم‌های آشوبگونه و کنترل پیش‌بین مدل.

- [37] S. Sabour, N. Frosst, and G. E. Hinton, "Dynamic routing between capsules," *Advances in neural information processing systems*, pp. 3856–3866, 2017.
- [38] R. Mukhometzianov and J. Carrillo, "CapsNet comparative performance evaluation for image classification," *arXiv preprint*, 2018.
- [39] F. Jafarzadehpour, A. S. Molahosseini, A. A. E. Zarandi, and L. Sousa, "Efficient Modular Adder Designs Based on Thermometer and One-Hot Coding," *IEEE Transactions on Very Large Scale Integration Systems*, vol. 27, no. 9, pp. 2142–2155, 2019.
- [40] N. Wajahat, S. Ahmed, A. Tahir, and H. Aqeel Khan. "Classification of breast cancer histology images using alexnet." In *International conference image analysis and recognition*, pp. 869–876. Springer, Cham, 2018.
- [41] K. M.K. "Contactless Palmprint Identification Using the Pretrained VGGNet Model." *Journal of Korea Multimedia Society*, vol. 21, no. 12, pp.1439–1447, 2018.
- [42] A. Qizwini, M., I. Barjasteh, H. Al-Qassab, and H. Radha. "Deep learning algorithm for autonomous driving using googlenet." *IEEE Intelligent Vehicles Symposium*. pp. 89–96, 2017.
- [43] The MNIST database of handwritten digits. Available online: <http://yann.lecun.com/exdb/mnist/>.
- [44] The CIFAR-10 dataset. Available online: <https://www.cs.toronto.edu/~kriz/cifar.html>.
- [45] fashion-mnist. Available online: <https://github.com/zalando-research/fashion-mnist>.