

ارائه یک روش مبتنی بر گرادیان و کرنل‌های شبکه عصبی عمیق به منظور تولید تصاویر بافتی

محمدحسین شکور

تغییر زاویه دید، تغییر اندازه، قرینه افقی یا عمودی، تغییر کنتراست و تغییر اندازه تصویر و غیره، از جمله مهم‌ترین راه‌های افزایش تعداد تصاویر در پردازش تصویر است [۱]. مشکل این روش‌ها آن است که تصاویر تولیدشده خیلی شبیه به تصاویر اولیه است و اغلب در شبکه‌های خیلی عمیق برای جلوگیری از بیش‌برازش کفایت نمی‌کند. دلیل آن این است که این داده‌ها تغییرات اساسی در تصاویر ایجاد نمی‌کنند، بلکه صرفاً برخی تغییرات بصری ساده در تصویر تولید می‌شود. یکی از مهم‌ترین انواع تصاویری که در پردازش تصویر و بینایی ماشین، چالش‌های زیادی دارد، تصاویر بافتی می‌باشد. تصاویر بافتی به دلیل ساختارهای پیچیده‌ای که دارند دارای چالش‌های خاصی برای پردازش تصویر هستند [۲].

همان گونه که اشاره شد یکی از مهم‌ترین چالش‌های هوش مصنوعی مسأله بیش‌برازش است [۳]. دلایل مختلفی برای این چالش وجود دارد و روش‌های مختلفی هم برای مقابله با آن ارائه شده است. در گذشته بیشتر بر این مسأله تأکید می‌شد که با ساده‌کردن شبکه و کاهش نرون‌های میانی و کاهش تعداد لایه‌های میانی شبکه، این مشکل مرتفع می‌شود، ولی این کاهش لایه‌های شبکه عصبی در حقیقت نوعی پاک‌کردن صورت مسأله بود؛ زیرا با کاهش تعداد لایه‌ها عملاً قدرت شبکه کمتر می‌شد و نمی‌توانست مسائل پیچیده‌تر را حل کند. از طرف دیگر با افزایش لایه‌ها و قدرتمندی شبکه مسأله بیش‌برازش به وجود می‌آمد؛ لذا روش‌هایی برای مقابله با این چالش ارائه شده است. برخی روش‌ها مثل کاهش تعداد مراحل آموزش، استفاده از داده‌های ارزیابی، حذف تصادفی نرون‌ها در حین آموزش، استفاده از ترم تنظیم‌کننده در تابع هزینه، زیاده‌کردن تعداد داده‌ها در هر دسته از داده‌های آموزش، استفاده از یادگیری انتقالی و شبکه‌های پیش‌آموزش‌داده‌شده و خیلی از طراحی‌های شبکه مثل لایه‌های ادغام در شبکه‌های کانولوشن از جمله این راه‌ها برای کاهش بیش‌برازش است [۳]. ولی همه این روش‌ها فقط تا حدی می‌توانند چالش بیش‌برازش را در یک شبکه عمیق مرتفع کنند. مهم‌ترین و اصلی‌ترین راه حل همان مسأله افزایش داده‌هاست؛ خصوصاً داده‌هایی که تنوع زیادی داشته باشند و فاقد نویز و داده‌های پرت باشند [۴]. شبکه‌های پیش‌آموزش‌داده‌شده نقش مهمی در گسترش و پیشرفت یادگیری عمیق داشته‌اند [۵]؛ زیرا با توجه به محدودیت‌های حافظه و پردازش در سیستم‌های فعلی، انجام بسیاری از پردازش‌ها روی داده‌های حجیم امکان‌پذیر نمی‌باشد و شبکه‌های پیش‌آموزش‌داده‌شده تا حد زیادی این مشکل را رفع کرده‌اند. در این مقاله از شبکه‌های پیش‌آموزش‌داده‌شده استفاده شده است؛ یعنی از ضرایب آماده شبکه‌های پیش‌آموزش‌داده‌شده برای تولید داده بافتی استفاده شده است. همه این شبکه‌های استفاده‌شده در این مقاله با داده ImageNet [۶] که شامل ۱۰۰۰ کلاس داده تصویری می‌باشد، آموزش داده شده‌اند. برخی از این شبکه‌ها عبارتند از:

چکیده: تولید پایگاه‌های داده تصاویر یکی از ضروریات بینایی ماشین به شمار می‌رود. روش‌های مختلفی مثل چرخش، تغییر روشنایی و زاویه دید، تغییر اندازه و ... برای افزایش داده‌های تصویری وجود دارد. عیب این روش‌ها آن است که تصاویر تولیدشده خیلی شبیه به تصاویر اولیه است و برای جلوگیری از بیش‌برازش کفایت نمی‌کند. شبکه‌های مواد تخصصی هم از همین مشکل برخوردار است. از بین همه انواع تصاویر، تصاویر بافتی چالش‌های بیشتری دارند؛ چون تصاویر بافتی را نمی‌توان مانند تصاویر معمولی با دوربین تصویربرداری تهیه کرد، زیرا بافت به راحتی در همه جا در دسترس نیست و هر نوع تصویری دارای بافت نیست. امروزه شبکه‌های عصبی عمیق به داده‌های حجیم برای آموزش نیاز دارند و در حال حاضر داده حجیم بافتی وجود ندارد. با استفاده از روش پیشنهادی می‌توان مجموعه داده‌های بافتی حجیم تهیه کرد. در این مقاله با استفاده از ضرایب کانولوشن شبکه‌های عمیق پیش‌آموزش‌داده‌شده بافت جدید تولید می‌شود. در این روش با اعمال گرادیان صعودی به تصاویر حاصل از فیلترهای کانولوشن، تصاویر بافتی جدید به طور مصنوعی تولید می‌شود. تفاوت این روش با روش‌های مولد آن است که این روش نیاز به تصاویر اولیه ندارد، بلکه در اینجا کلاس جدید بافتی تولید می‌گردد. پس از تولید بافت جدید، با روش‌های پردازش تصویر تعداد آن افزایش داده می‌شود. این روش بین ۳ تا ۵ برابر سریع‌تر از چند شبکه مولد مشهور است. کیفیت تصاویر هم خیلی بهتر است. یک نمونه پایگاه داده بافتی تولید شده که شامل ۲۴۰۰ تصویر در ۸۰ کلاس می‌باشد و در سایت Kaggle بارگذاری شده است.

کلیدواژه: تولید تصویر، تولید داده، گرادیان صعودی، شبکه عصبی کانولوشن.

۱- مقدمه

امروزه با طراحی شبکه‌های عصبی عمیق، نیاز به داده‌های حجیم بسیار بیشتر از گذشته حس می‌شود. تهیه داده متنوع یکی از چالش‌های مهم پردازش‌های شبکه‌های عمیق است. بدون داشتن داده کافی نه تنها یک شبکه عصبی عمیق نمی‌تواند گره‌گشا باشد، بلکه باعث مشکل بیش‌برازش^۱ و خطای زیاد در داده‌های تست می‌شود.

روش‌های مختلفی برای مقابله با بیش‌برازش وجود دارد که مهم‌ترین و اولین قدم افزایش داده^۲ است [۱]. در برخی از داده‌ها مانند داده‌های متنی، افزایش داده مشکل و پیچیده است. در تصاویر افزایش داده آسان‌تر انجام می‌شود. روش‌های رایج برای افزایش داده‌های تصویری مثل چرخش،

این مقاله در تاریخ ۲۸ خرداد ماه ۱۴۰۳ دریافت و در تاریخ ۶ مهر ماه ۱۴۰۳ بازنگری شد.

محمدحسین شکور (نویسنده مسئول)، دانشکده فنی و مهندسی، گروه مهندسی کامپیوتر، دانشگاه اراک، اراک، ایران، (email: mh-shakoor@araku.ac.ir).

1. Overfitting

2. Data Augmentation

جدول ۱: برخی روش‌های افزایش داده در شبکه‌های عمیق.

روش افزایش داده	مدل
جابه‌جایی، چرخش، تغییر روشنایی	AlexNet
چرخش، برش	ResNet
جابه‌جایی، چرخش، برش	DenseNet
زوم، جابه‌جایی، تغییر رنگ	YOLO

AlexNet [۷]، VGG۱۶ [۸]، ResNet [۹]، Xception [۱۰]، DenseNet [۱۱] و YOLO [۱۲]. در اغلب شبکه‌های عمیق از روش‌های افزایش داده برای زیاده‌تر کردن داده‌های تصویر استفاده شده است. در جدول ۱ لیستی از برخی از شبکه‌ها و روش‌هایی که برای افزایش داده استفاده کرده‌اند، نمایش داده شده است. این جدول برخی از شبکه‌های عمیق مشهور را نشان می‌دهد که از روش‌های مختلفی برای افزایش داده خود استفاده کرده‌اند.

به طور کلی روش‌های افزایش تصاویر به چهار دسته تقسیم می‌شوند: روش‌های فاقد مدل، روش‌های مبتنی بر مدل، روش‌های مبتنی بر بهینه‌سازی و روش‌های مبتنی بر پردازش تصویر. دسته اول یعنی روش‌های بدون مدل از مدل خاصی برای افزایش داده‌ها استفاده نمی‌کنند، در حالی که رویکرد مبتنی بر مدل از مدل‌های تولید تصویر برای ترکیب تصاویر استفاده می‌کند. در مقابل، روش‌های مبتنی بر بهینه‌سازی با هدف یافتن ترکیبی بهینه از تصاویر انجام می‌شود [۱۳]. روش‌های بر پایه پردازش تصویر هم کاربرد زیادی در ازدیاد داده‌های تصویری دارند.

در برخی از روش‌های افزایش داده تصویری، برچسب یا کلاس تصویر وجود ندارد؛ لذا این روش‌ها چالش تغییر کلاس داده را ندارند. ولی در برخی تصاویر دارای برچسب باید به مسأله کلاس تصویر دقت شود تا در اثر تغییرات اعمال‌شده در کلاس داده چالش ایجاد نشود. برخی از محققین مانند اینو [۱۴] با استفاده از دو تصویر متفاوت، یک کلاس تصویر جدید تولید کرده‌اند. در برخی مقالات مثل [۱۵] با ادغام خطی تصاویر، تصویر جدید ایجاد می‌شود. در برخی تحقیقات [۱۶] با استفاده از ۴ تصویر و ترکیب آنها تصویر جدید تولید می‌شود. هندریک و همکاران [۱۷] روشی ابداع کردند که ابتدا از یک تصویر، چندین تصویر تولید می‌کند و سپس از ترکیب آن تصاویر، تصویر جدید تولید می‌شود. روش‌هایی که از بهینه‌سازی استفاده می‌کنند نیز در تولید تصاویر نقش مهمی دارند. کیم و همکاران روشی مبتنی بر بهینه‌سازی ماسک ارائه کردند که دو تصویر را با استفاده از اطلاعات برجسته و آماری ترکیب می‌کند. در برخی روش‌ها برای تولید تصویر جدید از دو تصویر، ابتدا دو تصویر را به تکه‌های کوچک‌تر تقسیم کرده‌اند، سپس هر تکه را به‌صورت فضایی ترکیب می‌کنند و کناره‌های هر قسمت را به‌صورت خطی ادغام کرده‌اند [۱۸].

در برخی روش‌ها تصویر اشیاء را از پس‌زمینه جدا کرده و با قراردادن آنها در پس‌زمینه‌های دیگر، تصاویر جدید تولید می‌کنند [۱۹]. این تکنیک در اغلب مقالات به نام کپی-چسباندن یاد شده [۲۰] و کاربرد زیادی در تولید تصاویر جدید دارد؛ منتهی چالش اصلی آن جداسازی دقیق اشیاء از پس‌زمینه است. در ادامه این روش‌ها به‌صورت مجزا تشریح می‌شود.

در سال‌های اخیر مدل‌های مولد مبتنی بر شبکه‌های عصبی عمیق، نتایج قابل توجهی از نظر سنتز بافت تولید کرده‌اند. گودفلو و همکاران از

اولین افرادی هستند که شبکه‌های مولد متخاصم^۱ (GAN) را به این منظور معرفی کردند [۲۱]. شبکه‌های عصبی متخاصم شامل دو شبکه عصبی متمایزکننده و مولد هستند که نمونه‌های مختلفی از آنها در تولید داده‌های جدیدی معرفی شده‌اند. میرزا و اوسیندرو نسخه شرطی^۲ (CGAN) را معرفی کردند [۲۲] که می‌تواند با افزودن اطلاعات خارجی (برچسب‌ها) هم به تصاویر آموزشی و هم به تصاویر تولیدشده، تصاویر جدیدی تولید کند. در این تحقیق نشان داده شده که می‌توان خروجی مولد را با شرطی کردن مدل GANs بر روی اطلاعات کنترل کرد و شبکه CGAN به این منظور طراحی شده است. ژو و همکاران یک شبکه عصبی مولد با عنوان Cycle GAN را ارائه کردند [۲۳] که برای ترجمه تصویر به تصویر استفاده می‌شود. از ویژگی‌های این روش آن است که می‌تواند تصاویر را از یک دامنه به دامنه دیگر بدون نیاز به تناظر مستقیم تبدیل کند. برخی از روش‌ها با استفاده از انتخاب تکه‌هایی از بافت و کپی کردن آن در کنار بافت‌های دیگر، تصاویر بزرگ‌تر تولید می‌کنند [۲۴]. از محدودیت‌های این روش آن است که نمی‌تواند بافت‌های با ترکیبات جدید و پیچیده تولید کند و صرفاً مبتنی بر تکثیر همان بافت‌های قبلی است. برای ایجاد تنوع و ایجاد تغییرات بافتی برای رفع مشکل همسانی بافت‌ها در برخی از روش‌ها از تکنیک‌های مبتنی بر کوواریانس برای ایجاد و تحلیل ناهمسانی بافتی استفاده شده است [۲۵]. رادفورد و همکاران [۲۶] کلاسی از شبکه‌های عصبی کانولوشن را برای توسعه شبکه مولد معرفی کردند که هدف از آن، پایدار کردن آموزش شبکه نسبت به تغییرات پارامترهای آن است. جتچف و همکاران [۲۷] مدلی از یک شبکه عصبی مولد فضایی^۳ SGAN ارائه کردند که صرفاً با لایه‌های کانولوشن و بدون لایه‌های تماماً متصل با سرعت مناسب، عملیات تولید بافت را انجام می‌دهد. این روش اگرچه از سرعت مناسبی برخوردار است، در تشخیص کلاس بافت‌ها دچار مشکل می‌شود. از مشکلات محدودیت‌های مهم شبکه‌های مولد تخصصی برای تولید تصاویر این است که سرعت پردازش پایین است و تصاویر تولیدشده شباهت زیادی به تصاویر اولیه دارند. به همین منظور فن و همکاران با اضافه کردن لایه نرمال‌سازی دسته‌ای به شبکه VGG توانستند که سرعت تولید بافت را کمی افزایش دهند؛ ضمن اینکه با اضافه کردن نویز به تصاویر اولیه، تغییراتی در تصاویر خروجی نسبت به تصاویر اولیه ایجاد کردند [۲۸].

علاوه بر روش‌های مبتنی بر شبکه‌های مولد تخصصی، روش‌های قدیمی مانند مدل پنهان مارکف [۲۹] و نیز روش‌های غیرپارامتریک [۳۰] مختلفی هم برای تولید تصویر و بافت‌های تصویری ارائه شده‌اند که همه آنها نسبت به شبکه‌های عصبی عمیق از محدودیت‌ها و مشکلات بیشتری برخوردار هستند؛ لذا در اینجا مورد بحث زیادی قرار نگرفته‌اند. در این مقاله روشی مبتنی بر شبکه‌های عصبی عمیق ارائه شده که نسبت به روش‌های شبکه‌های عصبی متخاصم از مزیت‌های مختلفی برخوردار است. مزیت اول آن این است که برخلاف شبکه‌های عصبی در اینجا نیازی به تصاویر اولیه نیست. مزیت دیگر این است که تصاویر تولیدشده از کیفیت بیشتری برخوردارند و همچنین سرعت تولید آنها نسبت به شبکه‌های عصبی تخصصی بسیار بیشتر است.

سازماندهی این مقاله در ادامه بدین صورت است که در فصل ۲ انواع روش‌های افزایش داده‌های تصویری تشریح شده است. فصل ۳ و ۴ به

1. Generative Adversarial Network

2. Conditional GAN

3. Spatial GAN

برخی روش‌ها مبتنی بر یادگیری تقویتی هستند، مانند روش‌های BPA [۴۶]، Faster AA [۴۷]، Rand Augment [۴۸]، MADA [۴۹] و LDA [۵۰]؛ در حالی که روش‌هایی وجود دارد که از شبکه‌های متخاصم برای تولید تصاویر استفاده می‌کنند، مانند ADA [۵۱]، CDST-DA [۵۲] و AdaTransform [۵۳].

۲-۴ برخی روش‌های مبتنی بر پردازش تصویر

با توجه به اینکه یکی از مراحل روش پیشنهادی مبتنی بر پردازش تصویر است، در اینجا جزئیات بیشتری از این روش‌ها مرور می‌شود. به طور کلی روش‌های افزایش داده تصویری بر اساس پردازش تصویر را می‌توان به چند گروه زیر تقسیم کرد [۲]:

الف) قرینه‌سازی

استفاده از این نوع تغییر که یا به صورت آینه یا قرینه افقی است یا به صورت قرینه عمودی در اغلب تصاویر استفاده می‌شود. چرخش محور افقی بسیار رایج‌تر از چرخاندن محور عمودی است، زیرا ممکن است در اثر قرینه‌سازی عمودی، داده نادرست تولید شود. این روش یکی از ساده‌ترین روش‌های افزایش تصاویر است [۱].

ب) فضای رنگی

اعمال تغییرات در کانال‌های رنگی، روش دیگری برای تولید داده تصویری است. البته در پردازش تصویر، اغلب از داده‌های خاکستری استفاده می‌شود و پردازش رنگ کمتر در استخراج و مفهوم ویژگی دخالت داده می‌شود. تغییرات رنگی، ابتدا شامل جداسازی کانال‌های رنگی مانند R، G یا B می‌شود. علاوه بر این، مقادیر RGB را می‌توان به راحتی با عملیات ماتریسی ساده برای افزایش یا کاهش روشنایی تصویر تغییر داد. در اینجا هم باید توجه داشت که تغییرات رنگ با داده‌های واقعی تطابق داشته باشد و مثلاً برگ درخت نباید آبی‌رنگ شود [۱۳].

ج) برش

برش تصاویر را می‌توان به عنوان یک مرحله پردازش و تغییر تصویر برای افزایش داده‌های تصویر با ابعاد مختلف انجام داد. از برش تصادفی می‌توان برای تولید تصاویر جدید استفاده کرد. البته این تغییر در برخی موارد ممکن است اطلاعات مفید تصویر را حذف کند؛ لذا باید با احتیاط انجام شود. ضمن اینکه اغلب پایگاه‌های داده باید ابعاد تصاویر یکسان داشته باشند؛ لذا در اثر برش و کاهش تعداد نقاط برای حفظ ابعاد تصویر باید اندازه تصویر را بزرگ‌تر کرد یا نواحی خالی‌شده را با داده‌های قسمت‌های مجاور پر کرد [۹].

د) دوران

یکی از بهترین و ساده‌ترین روش‌های افزایش داده‌های تصویری همین روش است. این روش با چرخش تصویر بر روی محوری بین ۱ تا ۳۶۰ درجه انجام می‌شود. میزان چرخش باید توسط پارامتر درجه چرخش تعیین شود. چرخش‌های جزئی در محدود منفی تا مثبت مثلاً بیست درجه می‌تواند برای اغلب کارهای مفید باشد، اما با افزایش درجه چرخش، برچسب داده‌ها ممکن است حفظ نشود. لذا نکته مهم آن است که میزان چرخش نباید بیش از حد باشد؛ به طوری که داده نادرست تولید کند. مثلاً یک توپ را می‌توان به هر میزان چرخاند، ولی یک تصویر خوردو نهایتاً بیست یا سی درجه باید چرخیده شود؛ چون در واقعیت مثلاً خودروی عمودی وجود ندارد [۹].

ترتیب روش پیشنهادی و مقایسه نتایج پیاده‌سازی را نشان می‌دهد و فصل ۵ نتیجه‌گیری را بیان می‌کند.

۲- انواع روش‌های افزایش داده تصویری

این بخش بر انواع روش‌های افزایش داده تصویری مرور می‌کند.

۲-۱ روش‌های بدون مدل

روش‌های بدون استفاده از مدل از پردازش تصویر و تکنیک‌های آن برای افزایش داده‌های تصویری استفاده می‌کنند. این روش‌ها یا با یک تصویر و یا با چند تصویر، عملیات افزایش داده تصویری را انجام می‌دهند. در این روش‌ها با جابه‌جایی، تغییر رنگ و روشنایی و تغییرات هندسی، عملیات افزایش داده تصویری انجام می‌شود. برخی از روش‌ها با استفاده از یک تصویر، عملیات افزایش داده را انجام می‌دهند؛ مانند Hide-and-Seek [۳۱]، Cutout [۳۲]، Random Erasing [۳۳] و GridMask [۳۴]؛ ولی برخی دیگر با استفاده از چند تصویر عملیات تولید تصویر جدید را انجام می‌دهند. برخی از مهم‌ترین مقالات مرتبط با این روش‌ها عبارت هستند از: Sample Pairing [۳۵]، Mixup [۳۶]، BC Learning [۳۷]، CutMix [۳۸] و CutOut [۳۹].

۲-۲ روش‌های مبتنی بر مدل

این روش‌ها بر اساس یک مدل مولد و یک مدل متمایزکننده، عملیات تولید تصویر را انجام می‌دهند. تولید تصاویر با استفاده از این روش بسیار زمان‌بر است و نیاز به پردازش‌های سنگین دارد. این روش‌ها کلاً به سه دسته تقسیم می‌شوند [۱۳]: برخی روش‌ها بدون قید و شرط هستند، دسته دوم مشروط به برچسب یا کلاس هستند و دسته سوم مشروط به خود تصویر می‌باشند. اگر تصاویر فاقد کلاس یا برچسب داده باشند، در حقیقت تولید آنها هم بدون قید و شرط است؛ مانند CutPas [۴۰] و DCGAN [۴۱]. اما تصاویری که کلاس یا برچسب تصویر دارد باید با استفاده از روش‌های مشروط تولید شوند. در روش‌های مشروط هم یا برچسب تصویر بعد از افزایش داده حفظ می‌شود، مانند روش AugGAN [۴۲] و یا در برخی موارد، تغییرات طوری است که برچسب تصویر جدید عوض می‌شود، مانند روش‌های EmoGAN [۴۳] و GAN-MBD [۴۴].

۲-۳ روش‌های مبتنی بر بهینه‌سازی

روش‌های مبتنی بر بهینه‌سازی را می‌توان بر اساس یادگیری تقویتی یا شبکه‌های متخاصم انجام داد. شبکه‌های متخاصم حالت خاصی از شبکه‌های مولد هستند. همه الگوریتم‌های ذکرشده در دو روش قبلی، طرح‌های خاصی را نشان می‌دهند که در آنها، دانش اولیه‌ای از داده‌ها هست. در روش‌های مبتنی بر بهینه‌سازی، تغییر هاپرپارامتر چالش‌برانگیز و زمان‌بر است. یکی از راه‌های کاهش هزینه پردازشی، طراحی الگوریتم‌هایی است که استراتژی‌های بهینه‌سازی را تعیین می‌کنند. این الگوریتم‌ها دو دسته هستند: مبتنی بر یادگیری تقویتی و مبتنی بر یادگیری متخاصم. دسته اول از یادگیری تقویتی برای تعیین استراتژی بهینه استفاده می‌کند، در حالی که دسته دوم از شبکه‌های متخاصم مولد (GAN) استفاده می‌کنند. در حقیقت این شبکه‌ها هم می‌توانند برای تولید تصویر مبتنی بر مدل و هم مبتنی بر بهینه‌سازی تصویر استفاده شوند و هدف اتخاذشده در GANها تفاوت این دو گروه را رقم می‌زند. هدف روش‌های مبتنی بر مدل، تولید مستقیم تصاویر به جای اهداف دیگری مانند یافتن تبدیل‌های بهینه است، مانند روش ADA [۴۵].

جزئیات و لبه‌ها می‌شود، ولی در عوض نویز را کم می‌کنند. دسته دوم برخلاف نوع اول، هدفشان تیز کردن لبه‌هاست و البته باعث تشدید نویز می‌شوند. در شبکه‌های عمیق کانولوشن، این نوع فیلترها نقش بسیار مهمی در استخراج ویژگی از تصاویر ایفا می‌کنند. در این مقاله نیز از این فیلترها و با ترکیب آنها با گرادیان صعودی برای تولید تصاویر بافتی استفاده شده است. در روش‌های متداول، اعمال یک فیلتر یا ماسک کانولوشن به تصویر صرفاً تغییرات جزئی در لبه یا اطلاعات محلی تصویر ایجاد می‌کند و تصویر جدیدی خلق نمی‌کند. ولی در روش پیشنهادی کلاً یک تصویر بافتی جدید ایجاد می‌شود. این کار با استفاده از اعمال گرادیان به ماسک برحسب تصویر ورودی انجام می‌شود [۵۴].

ط) مخلوط کردن تصاویر

اختلاط تصاویر^۱ با هم با میانگین‌گیری مقادیر پیکسلشان، یک رویکرد نه‌چندان متداول برای افزایش داده است. تصاویری که با انجام این کار تولید می‌شوند، برای یک ناظر انسانی تبدیل مفیدی به نظر نمی‌رسند. یک عیب آشکار این تکنیک این است که از دیدگاه انسانی، منطقی نیست. درک یا توضیح افزایش عملکرد حاصل از مخلوط کردن تصاویر بسیار دشوار است؛ لذا استفاده از آن در همه جا متداول نیست [۱۵].

ی) پاک کردن تصادفی

این تکنیک به طور خاص برای مبارزه با چالش‌های تشخیص تصویر به دلیل انسداد طراحی شده است. انسداد به زمانی گفته می‌شود که برخی از قسمت‌های جسم نامشخص باشد. پاک کردن تصادفی با وادار کردن مدل به یادگیری ویژگی‌های توصیفی در مورد یک تصویر و جلوگیری از تطبیق بیش از حد آن به یک ویژگی بصری خاص در تصویر، انجام می‌شود. جدا از چالش بصری انسداد، پاک کردن تصادفی، یک تکنیک مناسب برای تضمین توجه شبکه به کل تصویر نه فقط به زیرمجموعه‌ای از آن است. این روش هم در بسیاری از تصاویر مرسوم نیست [۱۳].

۳- روش پیشنهادی

روش پیشنهادی از ۵ مرحله تشکیل شده و مرحله اصلی آن همان مرحله دوم است. کل مراحل در شکل ۱ آورده شده است. روش پیشنهادی ترکیبی از روش بهینه‌سازی گرادیان و روش مبتنی بر تولید تصویر مبتنی بر پردازش تصویر است. در این روش در مرحله ۱ و ۲ تصاویر بافتی ایجاد می‌شود و در مراحل بعد، عملیات افزایش تعداد تصاویر انجام می‌گردد. در برخی منابع [۵۴] از تکنیک گرادیان‌گیری برحسب ماسک‌های کانولوشن استفاده شده است. در این روش‌ها برای نمایش الگوهای استخراجی، هر کدام از ماسک‌های کانولوشن و همچنین تولید نقشه گرمایی تصاویر طبقه‌بندی‌شده به منظور نمایش قسمت‌های مهم تصویر برای طبقه‌بندی استفاده شده است [۵۵]. در این مقاله از چندین نوع شبکه عصبی پیش‌آموزش داده‌شده مانند VGG۱۶، VGG۱۹، Xception، DensNet و ResNet استفاده شده است.

این شبکه‌ها با داده بزرگ ImageNet [۶] آموزش داده شده‌اند و در اینجا قسمت کانولوشن آنها استفاده می‌شود. نحوه کار بدین صورت است که یک لایه از این شبکه‌ها انتخاب می‌شود. این لایه دارای تعدادی کانال یا همان ماسک کانولوشن است که هر بار یکی از این ماسک‌ها انتخاب می‌شود. ابتدا یک تصویر تصادفی تولید می‌شود، سپس این تصویر در یک حلقه تکرارشونده تغییر می‌کند. نحوه تغییر این تصویر به صورتی است

Step	Algorithm: The proposed Method
1	Using a Pre Trained Deep Neural Networks
2	Set α as learning rate Set Iterations Load a Pre Trained Deep Model For each Layer (L) of Model For each Convolution Mask (M) of Layer L Initialize an Image (Img) by Random Values For i=1 to Iterations Output = Convolution of Mask(M) to Img Loss = Average (Output) Img = GradientAscent (Loss, Img, α) End Save (Img of Mask (M) of Layer (L)) End End
3-1	For each Layer (L) of Model For each Image (Img) of each Mask (M) of Layer L Calculate LBP-M of Img Calculate the Percent of U2 (uniform pattern %) If ($U2 < T_L$) Remove Texture End End End
3-2	Remove Incomplete Textures Manually
4	Image Augmentation by Image Processing
5	Remove Marginal Areas and Save Central Area of Textures

شکل ۱: روش پیشنهادی.

ه) جابه‌جایی

جابه‌جایی تصاویر به چپ، راست، بالا یا پایین می‌تواند برای تولید تصاویر جدید مفید باشد. به عنوان مثال، اگر تمام تصاویر در یک مجموعه داده در مرکز قرار گیرند که در مجموعه داده‌های تشخیص چهره رایج است، می‌توان آن را با جابه‌جایی افزایش داد. همان طور که تصویر اصلی در یک جهت جابه‌جا می‌شود، فضای باقیمانده را می‌توان با یک مقدار ثابت مانند ۰ یا ۲۵۵ پر کرد یا می‌توان آن را با نویز تصادفی یا گاوسی پر کرد. نکته مهم آن است که در اثر جابه‌جایی ممکن است برخی اطلاعات نادرست در پس‌زمینه تصویر ایجاد شود که باید برای رفع آن اقدام کرد [۷].

و) تزریق نویز

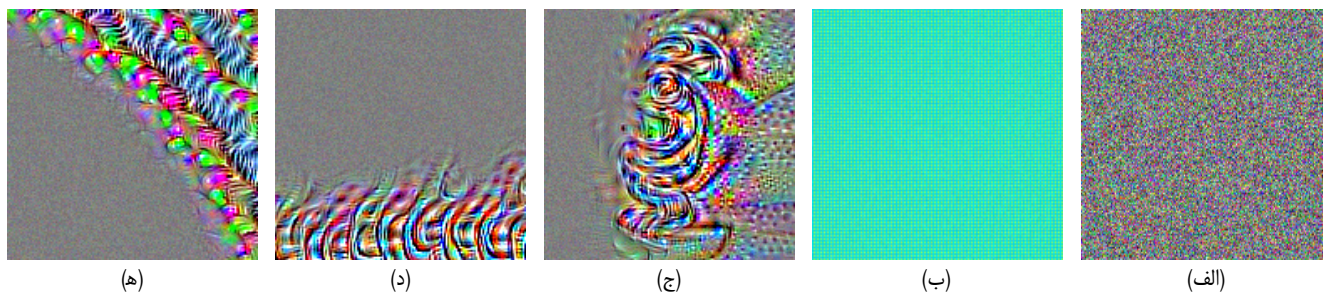
این روش با یک نویز ماتریسی، مقادیر تصادفی را که معمولاً از توزیع گاوسی گرفته می‌شود، به تصویر اضافه می‌کند. افزودن نویز به تصاویر می‌تواند به شبکه‌ها کمک کند تا ویژگی‌های قوی‌تری را بیاموزند. البته شدت نویز نباید از یک حد بیشتر باشد [۸].

ز) تغییرات بزرگ‌نمایی

تصاویر را اصطلاحاً می‌توان زوم کرد. این زوم کردن می‌تواند باعث کوچک‌تر یا بزرگ‌تر شدن تصویر شود. البته زوم در روش پردازش تصویر، صرفاً دیجیتال انجام می‌شود و می‌توان با روش‌های مختلف پردازش تصویر مثلاً با درون‌یابی و روش‌های پیچیده‌تر، کیفیت تصویر را بهبود داد [۱۲].

ح) فیلترهای کانولوشن

فیلترهای کانولوشن یک تکنیک بسیار محبوب در پردازش تصویر برای اعمال تغییرات در تصاویر هستند. این فیلترها به دو دسته پایین‌گذر و بالاگذر تقسیم می‌شوند که دسته اول باعث نرم شدن تصویر و حذف برخی



شکل ۲: برخی نمونه تصاویری که به صورت خودکار در مرحله ۳ حذف می‌شوند.

شکل ۲- الف است و هیچ بافتی تولید نکرده‌اند. در این مقاله یک روش خودکار در مرحله ۳-۱ برای تشخیص و حذف این تصاویر ارائه شده است. روش کار مرحله ۳-۱ بدین صورت است که ابتدا الگوی دودویی محلی [۵۶] از نوع اندازه برای بافت محاسبه می‌شود، سپس درصد الگوهای همگن بافتی محاسبه می‌شود. اگر این مقدار از یک حد کمتر باشد، آن تصویر مثل شکل ۲- الف فاقد بافت تشخیص داده شده و حذف می‌گردد. روابط (۱) تا (۳) نحوه محاسبه الگوی دودویی محلی اندازه را نمایش می‌دهد. در این مقاله از $P = 16$ و $R = 2$ برای این الگو استفاده شده است؛ زیرا نتایج بهتری نسبت به ناحیه‌های کوچک‌تر دارد.

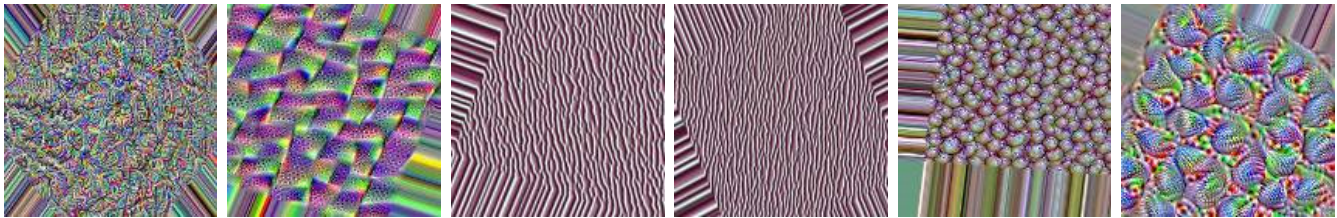
در شکل ۲- ب هم فقط رنگ بافت استخراج شده و مانند شکل ۲- الف به دلیل کم‌بودن درصد الگوهای همگن، فاقد بافت تشخیص داده شده [۵۶] و حذف می‌گردد. شکل‌های دیگر، تصاویر ناقص بافتی هستند که به دلایل مختلف از جمله کم‌بودن تعداد تکرار حلقه تولید بافت ناقص تولید شده‌اند. این تصاویر در مرحله ۳-۱ روش پیشنهادی به صورت خودکار شناسایی و حذف می‌شوند. شکل‌های ۲- ج تا ۲- ه بافت‌های ناقص هستند. در این نوع بافت‌ها با تقسیم بافت به ۴ یا ۹ ناحیه عملیات محاسبه LBP_M و شناسایی خودکار روی هر ناحیه انجام می‌شود و در صورتی که درصد الگوهای همگن حتی در یکی از نواحی کمتر از حد بود، آن بافت حذف می‌شود. قسمت اول مرحله ۳ یا همان مرحله ۳-۱ در واقع یک روش پیشنهادی برای شناسایی بافت‌های تولید نشده یا بافت‌های ناقص است. بدین منظور مطابق با شبه‌کد این قسمت، ابتدا باید الگوی دودویی محلی از نوع اندازه یعنی LBP_M برای شعاع $R = 2$ و تعداد نقاط الگو $P = 16$ برای تصویر محاسبه شود؛ سپس درصد الگوهای همگن یعنی درصد الگوهای با $U = 0$ و $U = 2$ محاسبه می‌گردد. مقدار U طبق (۳) محاسبه می‌شود. طبق [۴۶] این درصد برای بافت‌ها از یک حد بیشتر است. اگر این درصد از یک حد مثل T_L کمتر باشد، آن تصویر که مشابه تصاویر شکل ۲ است حذف می‌شود. هرچه لایه کانولوشن مورد استفاده به انتهای شبکه یعنی به لایه خروجی نزدیک‌تر باشد، مقدار این آستانه باید بزرگ‌تر شود و تعداد این تصاویر حذف شده بیشتر می‌شوند [۵۷]. آستانه T_L بر اساس میانگین دو تا از بزرگ‌ترین ویژگی‌های تولید شده توسط LBP_M یعنی میانگین دو مقدار ماکسیمم هیستوگرام LBP_M به دست می‌آید. در تصاویر بافتی ناقص، اولین (کد باینری همگی صفر) و آخرین مقدار (کد باینری همگی ۱) موجود در بردار ویژگی بزرگ‌تر از بقیه مقادیر هستند که علت آن عدم وجود اطلاعات بافتی کافی در تصویر است که باعث زیاد شدن الگوهای کاملاً صاف و فاقد بافت می‌شود.

در (۲) m_p میانگین کل اختلاف نقاط وسط و نقاط اطراف همه نواحی الگوی دودویی محلی است. اگر اندازه اختلاف هر نقطه میانی با نقطه اطراف از این مقدار بیشتر باشد، کد یک وگرنه کد صفر تولید می‌کند. سپس با اتصال این کدها و نگاشت آنها، ویژگی‌های بافتی به صورت

که گرادیان خروجی کانولوشن ماسک نسبت به تصویر محاسبه می‌شود و تابع هزینه، مجموع مقادیر خروجی ماسک کانولوشن انتخابی می‌باشد. با استفاده از گرادیان صعودی، این مقدار در یک حلقه با تغییر تصویر تصادفی اولیه زیادتر می‌شود. این تغییرات در نهایت منجر به تولید بافت جدید می‌شود. در اینجا یک حلقه با تکرار کافی روی داده‌های اولیه تصادفی اعمال می‌شود و با تغییر مقادیر نقاط این تصویر و با رابطه گرادیان صعودی، عملیات ماکسیمم کردن تصویر تولیدی حاصل از اعمال ماسک کانولوشن انتخابی به تصویر اولیه انجام می‌شود. در حقیقت در پایان این تکرارها تصویری تولید می‌شود که بیشترین گرادیان را بر اثر اعمال کانولوشن آن ماسک تولید می‌کند. به بیان دیگر تصویر تولید شده بیشترین حساسیت را نسبت به ماسک مربوطه دارد. در حقیقت این تصویر حاوی الگوهای بافتی از تصویرهای پایگاه ImageNet است که برای آموزش در شبکه‌های پیش‌آموزش داده شده مثل VGG16 استفاده شده است. در این مقاله از ماسک‌های انتخابی آموزش داده شده این شبکه برای تولید یک پایگاه داده بافتی استفاده شده است.

عملیات حلقه روی همه ماسک‌های لایه‌های مختلف کانولوشن اعمال می‌شود و تصاویر حاصل از این عملیات ذخیره می‌شود. در واقع در انتهای اعمال حلقه فوق، یک مجموعه از تصاویر وجود دارد که هر کدام مربوط به یک لایه کانولوشن است و هر کدام شامل تعدادی تصویر به تعداد کانال‌ها یا فیلترهای هر لایه می‌باشد. در اینجا فقط ۶۴ کانال از هر لایه انتخاب شده است. با بررسی تصاویر تولید شده مشخص می‌شود این تصاویر در حقیقت بافت‌های استخراج شده تصاویر توسط هر کدام از کانال‌های کانولوشن است. هرچه لایه کانولوشن به لایه ورودی نزدیک‌تر باشد، تصاویر تولیدی آن حاوی بافت‌های ریزتر و ساده‌تری است و تصاویر تولید لایه‌های آخر کانولوشن پیچیدگی بیشتری دارند. مطابق شبه‌کد شکل ۱ در مرحله دوم روش پیشنهادی، رابطه گرادیان صعودی از هر کدام از خروجی‌های حاصل از کانولوشن برحسب تصویر فعلی گرادیان صعودی محاسبه می‌شود.

مرحله سوم روش پیشنهادی شامل انتخاب تصاویر مناسب و حذف تصاویر ناقص است. به دلیل عملکرد برخی از کانال‌های کانولوشن لزوماً همه تصاویر تولید شده حاوی بافت‌های همسان نیست؛ ضمن اینکه برخی اصلاً حاوی بافت نیست و صرفاً رنگ یا اطلاعات ضعیف بافتی را تولید کرده‌اند. لذا در این مرحله برخی از تصاویر تولید مرحله قبل حذف می‌شود. اغلب در لایه‌های پایانی، احتمال اینکه برای یک تصویر ورودی، خروجی یک کانال فعال نشود زیادتر می‌شود و در حقیقت تصویر تصادفی اولیه با اعمال گرادیان تولید بافت نمی‌کند، زیرا ضرایب ماسک مربوطه تولید نشده‌اند. شکل ۲ برخی از این بافت‌های ناقص را نمایش می‌دهد. شکل ۲- الف خروجی این نوع ماسک‌های لایه‌ها است. مثلاً برای شبکه VGG16 در لایه block4_conv2 که جزء لایه‌های آخر کانولوشن محسوب می‌شود، از بین ۵۱۲ تصویر تولید شده، ۳۳۰ مورد آن مشابه



شکل ۳: برخی تصاویر نامناسب حاصل از مرحله ۴ افزایش داده.

جدول ۲: نمونه‌هایی از تصاویر تولیدی با استفاده از شبکه‌های مختلف.

مدل	Block ۱	Block ۲	Block ۳	Block ۴	Block ۵
VGG19					
ResNet101					
DensNet169					

که این تصاویر حاوی اطلاعات بافتی خاص‌تر و مربوط به کرنل‌های لایه‌های کانولوشن نهایی هستند. باید توجه داشت در الگوریتم پیشنهادی مرحله ۲، حلقه گرادیان صعودی ۱۰۰ بار تکرار می‌شود. با تغییر نرخ تکرار این حلقه، نتایج تصاویر بافتی هم تغییراتی جزئی خواهد داشت. هرچه حلقه تکرار بیشتری داشته باشد، تصویر بافتی تولیدشده شفاف‌تر و کامل‌تر است. البته تکرارهای زیاد تغییرات چندانی در تصویر ایجاد نمی‌کند؛ لذا در این مقاله برای تولید تصاویر تعداد تکرار ۱۰۰ انتخاب شده است. در پایان مرحله ۳ تصاویر بافتی تولیدشده مناسب برای تکثیر در مرحله بعد انتخاب می‌شوند. جدول ۲ نمونه‌هایی از این تصاویر بافتی تولیدشده را نشان می‌دهد.

در چهارمین مرحله از روش پیشنهادی با استفاده از پردازش تصویر برای افزایش داده مانند چرخش، تغییر نور، تغییر کنتراست، تقارن افقی، زوم کردن و غیره، تصاویر هر کدام از کلاس‌های بافتی تولیدی مرحله قبل تکثیر می‌شود. یک چالش این نوع افزایش داده آن است که مثلاً در اثر چرخش، جابه‌جایی یا زوم کردن (کوچک کردن) تصویر یک سری فضاهای خالی در کناره‌های تصویر به وجود می‌آید که روش‌های متداول افزایش داده، آن را به طور نامناسب پر می‌کند (مانند تصاویر شکل ۳). به منظور حذف این مشکلات در مرحله آخر، کناره‌های تصویر به اندازه ۵۰ پیکسل از هر طرف حذف می‌شود و ناحیه مرکزی تصویر به عنوان بافت نهایی انتخاب شده و در پایگاه داده قرار می‌گیرد. یعنی تصاویر مرحله قبل از اندازه ۲۵۰×۲۵۰ به ۱۵۰×۱۵۰ تبدیل گردیده و در پایگاه داده نهایی ذخیره می‌شوند.

۴- مقایسه نتایج

در اینجا یک پایگاه داده بافتی مصنوعی به نام Atex از ماسک‌های شبکه VGG16 تولید شده که شامل ۸۰ کلاس و ۳۰ تصویر در هر کلاس است.

برای تولید این پایگاه، ابتدا ۸۰ نوع تصویر جدید بافتی با روش پیشنهادی تولید شده است؛ یعنی ۸۰ کلاس بافتی جدید ایجاد شده است.

هیستوگرام تولید می‌شود [۵۷]. با توجه به آزمایش‌های انجام‌شده مشخص شد که برای تشخیص بهتر تصاویر بافتی مصنوعی تولیدشده از تصاویری مانند تصویر شکل ۲ که بافت تشکیل نداده‌اند یا ناقص هستند، روش LBP_MS^1 بهتر از LBP_S^2 عمل می‌کند. این مسأله خصوصاً برای لایه‌های اولیه که بافت‌های ریز تولید می‌کنند، بیشتر دیده می‌شود. مقدار $P=۱۶$ و $R=۲$ نیز بهتر از $P=۸$ و $R=۱$ عمل می‌کند

$$LBP_{P,R}^{riuV}(x,y) = \begin{cases} \sum_{i=1}^P m(g_i - g_c) & , U(LBP) \leq 2 \\ P+1 & , \text{otherwise} \end{cases} \quad (1)$$

$$m(g_i - g_c) = \begin{cases} 1 & , |g_i - g_c| \geq mg_c \\ 0 & , \text{otherwise} \end{cases} \quad (2)$$

$$U(LBP) = |m(g_{P-1} - g_c) - m(g_1 - g_c)| + \sum_{i=1}^P |m(g_i - g_c) - m(g_{i-1} - g_c)| \quad (3)$$

در قسمت دوم این مرحله یعنی ۲-۳ هم تصاویر ناقص بافتی به صورت دستی حذف می‌شوند. برخی از این تصاویر اغلب تصاویری با بافت‌های ریز و تکراری هستند، لذا خیلی از آنها باید حذف شود زیرا بافت جدیدی تولید نکرده‌اند. برخی دیگر از تصاویر، دو طیف بافتی را نتوانسته‌اند به طور کامل جدا کنند. البته می‌توان با تحقیقات بیشتری حذف این بافت‌ها را هم به صورت خودکار انجام داد و روشی را برای شناسایی آنها ارائه کرد.

در پایان مرحله سوم، تصاویر مناسب و حاوی اطلاعات بافتی مناسب و همگن انتخاب می‌شوند. برخی تصاویر حاوی اطلاعات بافتی ریزتر و عمومی‌تر هستند که از کانال‌های کانولوشن نزدیک به لایه ورودی تولید شده‌اند. در مقابل تصاویر حاصل از لایه‌های کانولوشن نزدیک به لایه خروجی از پیچیدگی بیشتری برخوردار هستند. دلیل آن هم این است

1. Magnitude Local Binary Pattern
2. Sign Local Binary Pattern

جدول ۳: تغییرات اعمال شده به منظور افزایش تصاویر در مرحله چهار.

نوع تغییر	تشریح
جابه‌جایی افقی	۲۰ درصد
جابه‌جایی عمودی	۲۰ درصد
چرخش	۳۰ درجه
برش	۲۰ درصد
زوم	۲۰ درصد
قرینه افقی	-
تغییر روشنایی	نصف تا ۱/۵ برابر

یک پایگاه داده حجیم بافتی برای پردازش توسط شبکه‌های عمیق ایجاد کرد. ابعاد تصاویر تولیدی هم کاملاً دلخواه است و می‌توان با زمان بیشتر، تصاویری با ابعاد بزرگ‌تر تولید کرد.

۴-۲ مقایسه بصری کیفیت تصاویر

یکی از روش‌های مقایسه نتایج روش پیشنهادی و روش‌های دیگر، مقایسه بصری خروجی آنهاست. جدول ۶ یک نمونه مقایسه بصری بین تصاویر بافتی تولیدشده در پایگاه ATeX و تصاویر تولیدشده بافتی با یک شبکه مولد [۶۵] را نمایش می‌دهد. همان گونه که مشاهده می‌شود، کیفیت تصاویر روش پیشنهادی خیلی بهتر از روش‌های مولد است. ضمن اینکه همان گونه که قبلاً اشاره شد، در روش پیشنهادی بافت جدید تولید می‌شود، ولی روش‌های مولد با استفاده از بافت‌های موجود تصاویر مشابه تولید می‌کنند.

۴-۳ مقایسه کمی نتایج

یک معیار مقایسه برای تصاویر تولیدشده بافتی همان جلوه‌های بصری و مقایسه آنها توسط کاربر به صورت چشمی است. در برخی مقالات از چندین معیار عددی برای اندازه‌گیری میزان وضوح و کیفیت تصاویر بافتی تولیدی استفاده شده است. در همه این مقالات تصاویر اولیه به شبکه مولد تخصصی داده می‌شود و تصاویر بافتی تولید می‌شود. برخی از این معیارها مثل معیار شباهت ساختاری $SSIM^1$ در شبکه‌های مولد استفاده شده [۶۶] که میزان شباهت تصاویر تولیدی با تصاویر اولیه را اندازه‌گیری می‌کند. اغلب این معیار را با میانگین مربعات خطا ترکیب می‌کنند. در برخی مقالات [۲۸] از معیار حداکثر نسبت سیگنال به نویز $PSNR^2$ هم استفاده شده است. برای محاسبه این دو معیار یعنی $SSIM$ و $PSNR$ علاوه بر تصاویر تولیدشده، نیاز به تصاویر اولیه هم است. معیارهای شباهت دیگری مثل معیار شباهت ساختاری بافت $STSIM^3$ نیز در برخی مقالات برای اندازه‌گیری شباهت تصاویر اولیه و تصاویر تولیدی استفاده شده‌اند [۶۷]. برخی معیارهای دیگر مثل میزان همبستگی بین تصاویر اولیه و تولیدی نیز وجود دارد که در همه آنها نیاز به تصاویر اولیه است. با توجه به اینکه در روش پیشنهادی ما تصاویر اولیه وجود ندارد، لذا باید از معیارهایی استفاده شود که فقط بر اساس تصاویر تولیدشده محاسبه می‌شوند و نیازی به تصاویر اولیه نباشد؛ مثل معیار G . این مقدار بر اساس میانگین گرادیان تصویر تولیدی محاسبه می‌شود و هرچه بالاتر باشد، نشان‌دهنده تصاویر با وضوح و جزئیات بهتر است [۲۸]. همچنین معیار انترپوی که میزان تغییرات و اطلاعات بافتی تصویر را نشان می‌دهد در اینجا برای مقایسه نتایج استفاده شده است.

جدول ۷ مقایسه نتایج روش پیشنهادی را با چند نمونه از بافت‌های مشهور نمایش می‌دهد. در این جدول دو ستون اول و دوم یعنی انترپوی و گرادیان معیارهایی هستند که هرچه بیشتر باشند، نشان‌دهنده کیفیت بهتر تصویر است، اما دو ستون سمت راست یعنی طبیعی بودن تصویر $NIQE^4$ و $BRISQUE^5$ معیارهایی با مفهوم منفی هستند و هرچه کمتر باشند، نشانه بهتری برای کیفیت تصویر هستند. این مقادیر میانگین مقادیر همه تصاویر هر پایگاه داده هستند. مطابق با این جدول بر اساس دو معیار اول

سپس با استفاده از روش‌های متداول پردازش تصویر که در جدول ۳ بیان شده، از هر تصویر بافتی ۲۹ تصویر تولید شده و در کل ۳۰ تصویر در هر کلاس ایجاد شده است؛ یعنی داده جدید مجموعاً شامل ۲۴۰۰ تصویر بافتی رنگی با اندازه 150×150 است. جدول ۴ چند نمونه از کلاس‌ها و تصاویر داده هر کدام را نمایش می‌دهد. اندازه فایل فشرده شده آن در حدود ۱۵۰ مگابایت است. می‌توان با این روش داده‌های بسیار بزرگ از بافت برای پردازش‌های عمیق و سنگین‌تر هم تولید کرد.

اغلب روش‌های اخیر تولید بافت مبتنی بر شبکه‌های تخصصی هستند. روش‌های قدیمی‌تر هم مبتنی بر تکنیک‌های پارامتریک هستند. در همه این روش‌ها با استفاده از بافت‌های موجود، تصاویر جدید تولید می‌شود که این تصاویر شباهت زیادی به بافت‌های اولیه دارد؛ لذا مسأله بیش‌برازش را نمی‌توانند به طور کامل حل کنند، زیرا تنوع بافتی و کلاس بافتی جدید تولید نمی‌کنند. از چالش‌های دیگر روش‌های موجود، زمان بسیار زیاد پردازش است. چالش سوم این است که تصاویر تولیدشده در اغلب روش‌ها نسبت به تصاویر طبیعی از اشکالات و ناهنجاری‌های زیادی برخوردار هستند. در این مقاله با استفاده از روش پیشنهادی، دو چالش اول تا حد زیادی و چالش سوم هم تا حد کمتری برطرف شده است. در ادامه نتایج روش پیشنهادی تشریح می‌گردد.

۴-۱ مقایسه با بزرگ‌ترین پایگاه‌های داده بافتی

جدول ۵ لیستی از مشهورترین و بزرگ‌ترین پایگاه‌های داده بافتی را نمایش می‌دهد [۵۸] تا [۶۴]. مشاهده می‌شود که بزرگ‌ترین پایگاه داده بافتی که در حال حاضر در مقالات علمی استفاده می‌شود، پایگاه داده ALOT است که ۲۷۵۰۰ تصویر بافتی در ۲۵۰ کلاس دارد. از لحاظ بیشترین تنوع کلاس نیز پایگاه STex با ۴۷۶ کلاس، بیشترین تعداد کلاس بافت را دارد. اکثر این داده‌ها فاقد رنگ و از نوع تک‌کانال یا تصاویر خاکستری هستند.

همان گونه که در جدول ۵ مشاهده می‌شود، تقریباً همه پایگاه‌های داده بافتی موجود از نوع داده حجیم محسوب نمی‌شوند و استفاده از آنها در شبکه‌های عمیق منجر به وقوع بیش‌برازش می‌گردد؛ لذا وجود یک پایگاه داده عمیق بافتی برای پردازش‌های عمیق ضروری به نظر می‌رسد. پایگاه داده تولیدشده در این مقاله یا ATeX صرفاً یک نمونه رنگی از پایگاه بافتی تولیدشده با روش پیشنهادی می‌باشد که در آن فقط از برخی کرنل‌های برخی لایه‌های شبکه VGG_{16} استفاده شده است. اگر همه کرنل‌ها و همه لایه‌های این شبکه استفاده شود، تعداد کلاس بیشتری می‌توان تولید کرد. با استفاده از کرنل شبکه‌های پیش‌آموزش داده‌شده دیگر نیز باز هم می‌توان کلاس داده بافتی را افزایش داد. همچنین در جدول ۳ با استفاده از روش‌های پردازش تصویر، تعداد هر تصویر ۳۰ برابر شده که به راحتی می‌توان این تعداد را به هر تعداد دلخواه افزایش داد و

1. Structural Similarity Index Measure
2. Peak Signal to Noise Ratio
3. Structural Texture Similarity Index Measure
4. Naturalness Image Quality Evaluator
5. Blind/Refereceless Image Spatial Quality Evaluator

جدول ۴: چند نمونه از بافت‌های مصنوعی ATEX.

کلاس	نمونه تصاویر			
۵				
۲۰				
۴۵				
۶۶				
۷۵				

جدول ۶: مقایسه تصاویر تولیدی روش پیشنهادی و شبکه PGGAN.

شبکه PGGAN	روش پیشنهادی

جدول ۵: برخی از مشهورترین و بزرگ‌ترین پایگاه‌های داده بافتی.

نوع تصویر	ابعاد تصویر	کل نمونه‌ها	نمونه هر کلاس	تعداد کلاس	داده
رنگی	۲۰۰×۲۰۰	۵۶۱۲	۹۲	۶۱	CURet [۵۸]
خاکستری	۱۲۸×۱۲۸	۴۳۲۰	۱۸۰	۲۴	Outex۱۰ [۵۹]
خاکستری	۱۲۸×۱۲۸	۹۱۲۰	۳۸۰	۲۴	Outex۱۲ [۵۹]
خاکستری	۱۲۸×۱۲۸	۱۳۶۰	۲۰	۶۸	Outex۱۳ [۵۹]
خاکستری	۱۲۸×۱۲۸	۷۶۱۶	۱۶	۴۷۶	STex [۶۰]
خاکستری	۶۴۰×۴۸۰	۱۰۰	۴۰	۲۵	UIUC [۶۱]
خاکستری	۱۲۸×۱۲۸	۱۷۹۲	۱۶	۱۱۲	Brodatz [۶۲]
خاکستری	۱۲۸×۱۲۸	۱۱۰۴	۱۶	۶۹	MeasTex [۶۳]
خاکستری	۵۱۲×۷۶۸	۲۷۵۰۰	۱۱۰	۲۵۰	A LOT [۶۴]
رنگی	۱۵۰×۱۵۰	۲۴۰۰	۳۰	۸۰	ATex.VGG۱۶

جدول ۷: مقایسه بافت‌های تولیدی با بافت‌های مشهور.

داده	معیارهای مثبت		معیارهای منفی	
	Entropy	Gradient	NIQE	BRISQUE
CURet [۵۸]	۵,۵۷	۶۹,۷۸	۱۹,۰۷	۳۹,۲۸
UIUC [۶۱]	۶,۸۳	۶۹,۳۵	۶,۱۴	۳۵,۸۸
OUTEX [۵۹]	۵,۱۳	۵۷,۶۶	۱۸,۸۸	۴۱,۰۰
پیشنهادی	۷,۶۴	۱۷۰,۸۳	۱۸,۸۸	۳۸,۵۷

برای مقایسه بافت‌های تولیدی با بافت‌های تولیدشده با روش‌های دیگر از جمله شبکه‌های GAN جدول ۸ ارائه شده است. در این جدول معیارهای محاسبه‌شده برای پایگاه داده تولیدی با نتایج تصاویر بافتی تولیدشده با ۵ نوع شبکه تخصصی [۲۸]، [۶۵]، [۶۷] و [۶۸] و یک روش پارامتریک [۶۹] مقایسه شده‌اند. آخرین ردیف این جدول مربوط به روش پارامتریک است که بر اساس این نتایج، عملکرد این روش بر اساس دو معیار منفی بهتر از همه شبکه‌های GAN و روش پیشنهادی است؛ ولی

بافت‌های تولیدشده با روش پیشنهادی از وضوح و کیفیت بهتری نسبت به بافت‌های معروف موجود در جدول برخوردار هستند. بر اساس دو معیار منفی موجود در جدول، تصاویر بافتی UIUC بهتر از بقیه هستند؛ زیرا این تصاویر شامل بافت‌های بسیار واضح و با جزئیات دقیق و زیاد هستند. بر اساس معیار میزان طبیعی‌بودن بافت یعنی معیار NIQU، تصاویر بافتی تولیدی مشابه تصاویر OUTEX و CURet هستند. همچنین بر اساس معیار BRISQUE نیز تصاویر تولیدی کمی بهتر از بافت‌های OUTEX و CURet هستند.

جدول ۱۰: نتایج الگوی دودویی محلی روی داده ATEX.

	$R=1, P=8$	$R=2, P=16$	$R=3, P=24$
LBP_S _{P,R} ^{riu}	۴۶/۱۷	۷۸/۴۴	۸۳/۱۹
LBP_M _{P,R} ^{riu}	۶۱/۳۶	۸۰/۴۲	۸۲/۴۴
LBP_SM _{P,R} ^{riu}	۷۷/۹۷	۹۲/۸۹	۹۲/۵۳
LBP_SMC _{P,R} ^{riu}	۸۴/۸۳	۹۴/۰۶	۹۴/۱۴

شامل ۱۳ GB حافظه همراه با دو پردازنده زنون ۲/۲ GHz مجازی است. در روش PGGAN از یک سخت‌افزار قدرتمند شامل دو پردازنده زنون ۲.۰ GHz ۲۰C ۶۱۳۸ همراه با ۱۹۲ GB حافظه همراه با دو پردازنده گرافیکی NVIDIA Tesla V1۰۰ استفاده شده است. در این روش زمان ۵ روز برای تولید ۱۰۰۰ تصویر با ابعاد ۱۰۲۴×۱۰۲۴ صرف شده است [۶۵]؛ یعنی به طور متوسط ۴۳۲ ثانیه برای تولید هر تصویر بافتی صرف شده است. نمونه‌ای از این تصاویر در ستون سمت چپ جدول ۶ نمایش داده شده است.

برای مقایسه عادلانه، زمان پردازش روش پیشنهادی برای تولید تصاویر با ابعاد ۲۵۶×۲۵۶ و ۱۰۲۴×۱۰۲۴ هم محاسبه و در جدول ۹ لحاظ شده است. روش پیشنهادی برای تولید تصاویر اندازه تصاویر PGGAN یعنی تصاویر با ابعاد ۱۰۲۴×۱۰۲۴ به طور متوسط ۸۵ ثانیه زمان لازم دارد که تقریباً ۵ برابر سریع‌تر از PGGAN است. ضمن اینکه سخت‌افزارها و پردازنده مورد استفاده در روش PGGAN خیلی قدرتمندتر از گوگل کولب رایگان است. در جدول ۹ زمان دو روش مشهور دیگر هم برای تولید بافت نمایش داده شده است. دو روش DCGAN و SAGAN [۶۸] و [۷۰] طبق این جدول به ترتیب ۴۵/۵ و ۲۷/۷ ثانیه زمان برای تولید بافت‌هایی به ابعاد ۲۵۶×۲۵۶ نیاز دارند که این زمان برای روش پیشنهادی ۸/۹ می‌باشد؛ یعنی به ترتیب تقریباً ۵/۱ و ۳/۱ برابر روش پیشنهادی است. البته سخت‌افزار مورد استفاده در این دو روش [۶۸] نیز قدرتمندتر از گوگل کولب رایگان است.

۴-۵ طبقه‌بندی پایگاه داده

یک نمونه پایگاه داده بافتی مصنوعی با روش پیشنهادی تولید شده که حاوی بافت‌های مصنوعی تولیدشده در این مقاله است و همراه با کد متلب طبقه‌بندی آن توسط چند الگوی محلی دودویی در سایت kaggle در قسمت datasets با نام ATex-dataset قرار داده شده است. جدول ۱۰ نتایج طبقه‌بندی این بافت را با چند الگوی دودویی محلی با اندازه ناحیه محلی مختلف نشان می‌دهد. در این طبقه‌بندی از اطلاعات رنگی بافت‌ها استفاده نشده و تصاویر ابتدا تبدیل به خاکستری شده‌اند. در این نتایج نیمی از تصاویر هر کلاس به صورت تصادفی برای آموزش و نیم دیگر برای تست استفاده شده و میانگین نتایج بعد از ۱۰ بار تکرار در جداول نمایش داده شده است. روش LBP_SMC همان الگوی دودویی محلی کامل CLBP است [۷۱] که بهترین نتایج را در اغلب داده‌ها تولید کرده و در اغلب پایگاه‌های بافتی، نتیجه طبقه‌بندی با این روش تا نزدیک ۱۰۰ درصد رسیده است [۵۷] و [۷۱].

در اینجا بهترین نتیجه تولیدی هنوز ۶ درصد تا دقت کامل فاصله دارد؛ لذا این میزان خطا از این داده را می‌توان برای به چالش کشیدن توصیفگرهای تصاویر بافتی استفاده کرد. ضمن اینکه در اینجا از نیمی از داده‌ها برای تست و از نیمی برای آموزش استفاده شده است. اگر این نسبت تغییر کند و مثلاً ۱۰ درصد داده برای تست استفاده شود، طبیعتاً دقت طبقه‌بندی کمتر می‌شود و چالش بافتی بیشتری ایجاد می‌شود.

جدول ۸: مقایسه بافت‌های تولیدی با روش‌های دیگر تولید بافت.

روش	معیارهای مثبت		معیارهای منفی	
	Entropy	Gradient	NIQE	BRISQUE
Fan, et al. [۲۸]	۶/۵۳	۷۵/۵۰	۶/۶۶	۴۱/۸۳
StyleGAN۳ [۶۷]	۶/۴۳	۵۱/۵۷	۸/۸۳	۴۳/۶۵
PGGAN [۶۵]	۵/۹۷	۱۳/۳۴	۸/۶۸	۴۹/۴۴
DCGAN [۶۸]	۵/۵۴	۳۳/۳۴	۵/۵۱	۵۱/۹۳
SAGAN [۶۸]	۵/۷۸	۳۳/۶۸	۵/۳۶	۵۱/۷۹
Portilla & Simonelli [۶۹]	۷/۶۸	۶۷/۸۵	۳/۸۸	۲۷/۳۷
پیشنهادی	۷/۶۴	۱۷۰/۸۳	۱۸/۸۸	۳۸/۵۷

جدول ۹: مقایسه متوسط زمان تولید هر تصویر بافتی جدید.

روش	ابعاد تصویر	زمان تولید یک تصویر (sec)
روش پیشنهادی	۱۵۰×۱۵۰	۳/۳
روش پیشنهادی	۲۵۶×۲۵۶	۸/۹
روش پیشنهادی	۱۰۲۴×۱۰۲۴	۸۵
PGGAN [۶۵]	۱۰۲۴×۱۰۲۴	۴۳۲
DCGAN [۶۸]	۲۵۶×۲۵۶	۴۵/۵
SAGAN [۶۸]	۲۵۶×۲۵۶	۲۷/۷

بر اساس دو معیار انترپوی و گرادیان روش پیشنهادی از همه روش‌های دیگر عملکرد بهتری داشته است. به معنای دیگر روش پارامتریک بافت‌های طبیعی‌تری نسبت به شبکه‌های تخصصی و روش پیشنهادی تولید کرده است؛ ولی از لحاظ اطلاعات وجود در بافت و جزئیات و پیچیدگی بافتی، روش پیشنهادی خیلی عملکرد بهتری نسبت به سایر روش‌ها داشته است.

۴-۶ مقایسه زمان پردازش

یکی از چالش‌های مهم پردازش در شبکه‌های عمیق، مسأله زمان آموزش و محدودیت‌های حافظه و پردازنده است. روش‌های مختلفی برای غلبه بر این محدودیت وجود دارد که در این مقاله از یکی از آنها استفاده شده است. استفاده از روش‌های مبتنی بر شبکه‌های پیش‌آموزش داده‌شده یکی از روش‌هایی است که این چالش را تا حد زیادی برطرف می‌کند. در روش پیشنهادی، فاز آموزش مشابه آنچه که در شبکه‌های مولد است به نوعی وجود ندارد؛ زیرا با استفاده از ضرایب کرنل‌های لایه‌های کانولوشن یک شبکه پیش‌آموزش داده‌شده عملیات تولید بافت جدید انجام می‌شود. قسمت اصلی زمان پردازش روش پیشنهادی همان حلقه مرحله دوم در شکل ۱ است که برای تولید تصاویر اولیه مرحله دوم با ابعاد ۳۰۰×۳۰۰ و ۱۰۰ بار تکرار حلقه به‌ازای هر تصویر حدود ۲/۳ ثانیه زمان لازم است. سپس طی مرحله بعد، تصاویر بافتی ناقص و معیوب حذف می‌شود و حاشیه‌های اطراف تصاویر تولیدی بعد از افزایش داده، برش زده می‌شود تا ابعاد ۱۵۰×۱۵۰ تولید شود. به ازای هر تصویر دو مرحله بعدی روش پیشنهادی نیز مجموعاً کمتر از ۱ ثانیه زمان لازم دارد؛ لذا متوسط زمان برای تولید یک بافت در حدود ۳/۳ ثانیه است.

جدول ۹ مقایسه‌ای بین زمان لازم برای تولید یک تصویر بافتی با روش پیشنهادی و زمان متناظر در چند شبکه مولد را نمایش می‌دهد. یکی از این روش‌های مولد برای تولید بافت به نام [۶۵] می‌باشد. همان گونه که مشاهده می‌شود، زمان متوسط پردازش لازم برای یک تصویر با روش پیشنهادی بسیار کمتر از این روش مولد می‌باشد. در روش پیشنهادی از گوگل کولب ورژن رایگان برای پردازش استفاده شده که

۵- نتیجه گیری

کیفیت بافت‌های تولیدشده و استفاده از آن در کاربردهای عملی مانند تولید نقشه‌های بافتی است. از مهم‌ترین پیشنهادهایی که می‌توان برای تحقیقات آینده در این مقاله ارائه کرد، تولید نقشه‌های بافتی مثلاً برای چاپ در پارچه، فرش، موکت و کلاً صنایع بافندگی است. از آنجا که تولید این بافت‌ها مصنوعی و نامحدود می‌باشد، می‌توان انواع و اقسام جدیدی از طرح‌ها و نقش‌های بافتی برای صنایع بافندگی و چاپ تولید و استفاده نمود. همچنین مرحله حذف دستی بافت‌ها هم می‌تواند با ارائه روشی به‌صورت خودکار انجام شود.

مراجع

- [1] C. Shorten and M. T. Khoshgoftaar, "A survey on image data augmentation for deep learning," *Mathematics and Computers in Simulation*, vol. 6, no. 60, pp. 1-48, Jul. 2019.
- [2] S. Di Cataldo and E. Ficarra, "Mining textural knowledge in biological images: applications, methods and trends," *Computational and Structural Biotechnology J.*, vol. 15, pp. 56-67, 2017.
- [3] S. Nitish, H. Geoffrey, K. Alex, S. Ilya, and S. Ruslan, "Dropout: a simple way to prevent neural networks from overfitting," *J. Mach Learn Res.*, vol. 15, no. 1, pp. 1929-1958, 2014.
- [4] P. Luis and W. Jason, "The Effectiveness of Data Augmentation in Image Classification Using Deep Learning," Stanford University Research Report, 2017.
- [5] D. Erhan, Y. Bengio, A. Courville, P. A. Manzagol, and P. Vincent, "Why does unsupervised pre-training help deep learning?" *J. Mach. Learn. Res.*, vol. 11, pp. 625-660, 2010.
- [6] J. Deng, et al., "ImageNet: a large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, pp. 248-255, Miami, FL, USA, 20-25 Jun. 2009.
- [7] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Adv Neural Inf Process Syst.*, vol. 25, no. 1, pp. 106-114, 2012.
- [8] S. Karen and Z. Andrew, "Very Deep Convolutional Networks for Large-Scale Image Recognition," arXiv preprint arXiv:1409.1556, 2014.
- [9] H. Kaiming, Z. Xiangyu, R. Shaoqing, and S. Jian, "Deep residual learning for image recognition," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 770-778, Las Vegas, NV, USA, 27-30, 2016.
- [10] S. Christian, V. Vincent, I. Sergey, S. Jon, and W. Zbigniew, "Rethinking the Inception Architecture for Computer Vision," arXiv e-prints arXiv:1512.00567, 2015.
- [11] H. Gao, L. Zhuang, M. Laurens, and Q. W. Kilian, "Densely Connected Convolutional Networks," arXiv preprint arXiv:1608.06993, 2016.
- [12] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: unified, real-time object detection," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 779-788, Las Vegas, NV, USA, 27-30 Jun. 2016.
- [13] M. Xu, S. Yoon, A. Fuentes, and D. S. Park, "A comprehensive survey of image augmentation techniques for deep learning," *Pattern Recognition*, vol. 137, Article ID: 109347, 2023.
- [14] H. Inoue, "Data Augmentation by Pairing Samples for Images Classification," arXiv preprint arXiv:1801.02929, 2018.
- [15] H. Zhang, M. Cisse, Y. N. Dauphin, and D. Lopez-Paz, "Mixup: Beyond Empirical Risk Minimization," arXiv preprint arXiv:1710.09412, 2017.
- [16] A. Bochkovskiy, C. Y. Wang, and H. Y. M. Liao, "Yolov4: Optimal Speed and Accuracy of Object Detection," arXiv preprint arXiv:2004.10934, 2020.
- [17] D. Hendrycks, et al., "Augmix: a simple method to improve robustness and uncertainty under data shift," in *Proc. the 18th Int. Conf. on Learning Representations*, 6 pp. 6, 26 Apr.-1 May 2020.
- [18] K. Baek, D. Bang, and H. Shim, "Gridmix: strong regularization through local context mapping," *Pattern Recognit.*, vol. 109, Article ID: 107594, 2021.
- [19] D. Dwibedi, I. Misra, and M. Hebert, "Cut, paste and learn: surprisingly easy synthesis for instance detection," in *Proc. of the IEEE Int. Conf. on Computer Vision*, pp. 1301-1310, Venice, Italy, 22-29 Oct. 2017.
- [20] H. Naveed, S. Anwar, M. Hayat, K. Javed, and A. Mian, "Survey: image mixing and deleting for data augmentation," *Engineering Applications of Artificial Intelligence*, vol. 131, no. C, Article ID: 107791, 2024.

در این تحقیق یک روش سریع و مبتنی بر استفاده از فیلترهای کانولوشن تولیدشده از شبکه‌های پیش‌آموزش‌دیده ارائه شد که هدف آن تولید تصاویر بافتی جدید است. در این روش در چند مرحله، تصاویر بافتی با استفاده از گرادایان صعودی و میزان حساسیت هر فیلتر به الگوهای مختلف تصویری تولید می‌شود و سپس پس از حذف تصاویر ناقص از بین تصاویر مناسب تولیدشده با روش‌های متداول داده‌افزایی تصویری زیاد می‌شوند و در کلاس‌های مختلف بافتی قرار می‌گیرند. روش پیشنهادی برخلاف شبکه‌های مولد نیاز به تصویرهای اولیه بافتی ندارد، زیرا از همان ضرایب محاسبه‌شده کانولوشن در شبکه پیش‌آموزش‌دیده‌شده استفاده می‌کند. همچنین این روش برخلاف روش‌های شبکه مولد عمیق، سرعت بسیار زیادتری دارد و نیاز به پردازش‌های سنگین ندارد.

از مهم‌ترین مزیت‌های روش پیشنهادی نسبت به روش‌های متداول، تولید بافت جدید است، زیرا در روش‌های داده‌افزایی متداول بافت جدیدی یا کلاس بافتی جدید تولید نمی‌شود؛ بلکه به همان بافت‌های اولیه تغییرات بصری اعمال و تصاویر تکثیر می‌شود. با روش پیشنهادی می‌توان با استفاده از شبکه‌های پیش‌آموزش‌دیده مختلف پایگاه‌های داده بافتی بزرگ تولید کرد. در حقیقت مهم‌ترین ویژگی روش پیشنهادی، امکان تولید پایگاه‌های داده حجیم بافتی با برای پردازش‌های عمیق است؛ زیرا امکان تهیه این پایگاه‌های بافتی به‌صورت طبیعی و مشابه با تصاویر دیگر وجود ندارد، زیرا تعداد و تنوع بافت‌های طبیعی با محدودیت همراه است. در حالی که روش پیشنهادی محدودیتی در تولید تعداد بافت ندارد. همچنین در مقایسه با شبکه‌های عصبی مولد، روش پیشنهادی از مزیت زیادی برخوردار است، زیرا اولاً سرعت تولید تصاویر در آن بسیار بیشتر از شبکه‌های مولد است. ثانیاً کیفیت تصاویر نسبت به شبکه‌های مولد بهتر است و دلیل سوم و مهم‌تر آن است که در شبکه‌های مولد تصویر تولیدشده شباهت زیادی با تصاویر اولیه دارد؛ در حالی که در روش پیشنهادی اصلاً تصویر اولیه‌ای از بافت وجود ندارد و این تصاویر تولیدشده در حقیقت از داخل تصاویر متداول استخراج شده است.

نتایج کمی مقایسه تصاویر تولیدشده نشان می‌دهند روش پیشنهادی از لحاظ تولید بافت‌های پیچیده و دارای جزئیات بافتی از داده‌های مشهور بافتی نتایج بهتری تولید کرده است. اگرچه از لحاظ طبیعی بودن نسبت به برخی بافت‌ها کیفیت کمتری دارد، ولی با برخی بافت‌های دیگر نتایج مشابه دارد. ضمن اینکه مقایسه روش پیشنهادی با روش‌های تولید بافت تخصصی نشان می‌دهد روش پیشنهادی هم از لحاظ پیچیدگی و جزئیات بافتی و هم از لحاظ طبیعی بودن نتایج نسبت به شبکه‌های تخصصی بهتر عمل کرده است؛ اگرچه نسبت به روش‌های پارامتریک، تصویری با حالت طبیعی کمتری تولید کرده است.

یکی از پیشنهادهایی که می‌توان برای تولید پایگاه بافتی بهتر ارائه کرد این است که در روش پیشنهادی، ابتدا یک تصویر از یک کلاس بافتی جدید تولید شده و سپس با روش‌های پردازش تصویر این تصویر تکثیر می‌شود. این کار ممکن است باعث نشستی داده^۱ در حین آموزش و یادگیری ماشین گردد؛ لذا پیشنهاد می‌شود به‌جای استفاده از یک تصویر تولیدی و تکثیر آن، تعداد بیشتری تصویر بافتی جدید از یک کلاس استفاده شود تا این مشکل هم مرتفع گردد.

یکی از کارهای مهمی که می‌توان در ادامه این مقاله انجام داد، بهبود

- [46] D. Ho, E. Liang, X. Chen, I. Stoica, and P. Abbeel, "Population based augmentation: efficient learning of augmentation policy schedules," in *Proc. 36th Int. Conf. on Machine Learning*, pp. 2731-2741, Long Beach, CA, USA, 9-15 Jun. 2019.
- [47] R. Hataya, J. Zdenek, K. Yoshizoe, and H. Nakayama, "Faster autoaugment: learning augmentation strategies using backpropagation," in *Proc. European Conf. on Computer Vision*, pp. 1-16, Munich, Germany, 23-28 Aug. 2020.
- [48] E. D. Cubuk, B. Zoph, J. Shlens, and Q. V. Le, "RandAugment: practical automated data augmentation with a reduced search space," in *Proc. of the IEEE/CVF Conf. on Computer Vision and Pattern Recognition Workshops*, pp. 702-703, Seattle, WA, USA, 14-19 Jun. 2020.
- [49] R. Hataya, J. Zdenek, K. Yoshizoe, and H. Nakayama, "Meta approach to data augmentation optimization," in *Proc. of the IEEE/CVF Winter Conf. on Applications of Computer Vision*, pp. 2574-2583, Waikoloa, HI, USA, 3-8 Jan. 2022.
- [50] B. Zoph, et al., "Learning data augmentation strategies for object detection," in *Proc. European Conf. on Computer Vision*, pp. 566-583, Munich, Germany, 23-28 Aug. 2020.
- [51] P. Li, X. Liu, and X. Xie, "Learning sample-specific policies for sequential image augmentation," in *Proc. of the 29th ACM Int. Conf. on Multimedia*, pp. 4491-4500, 20-24 Oct. 2021.
- [52] A. J. Ratner, H. Ehrenberg, Z. Hussain, J. Dunnmon, and C. Ré, "Learning to compose domain-specific transformations for data augmentation," in *Proc. of the 31st Int. Conf. on Neural Information Processing Systems*, pp. 3239-3249, Long Beach, CA, USA, 4-9 Dec. 2017.
- [53] Z. Tang, X. Peng, T. Li, Y. Zhu, and D. N. Metaxas, "Adatransform: adaptive data transformation," in *Proc. of the IEEE/CVF Int. Conf. on Computer Vision*, pp. 2998-3006, Seoul, South Korea, 27 Oct.-2 Nov. 2019.
- [54] F. Chollet, *Deep Learning with Python*, Manning Publication, 2017.
- [55] R. R. Selvaraju, et al., Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization, arXiv preprint arXiv:1610.02391, 2017.
- [56] F. Bianconi and A. Fernandez, "On the occurrence probability of local binary patterns: a theoretical study," *J. Math. Imag. Vis.*, vol. 40, no. 3, pp. 259-268, 2011.
- [57] M. H. Shakoar and R. Boostani, "Extended mapping local binary pattern operator for texture classification," *International J. of Pattern Recognition and Artificial Intelligence*, vol. 3, no. 6, Article ID: 1750019, 2017.
- [58] K. J. Dana, B. Van Ginneken, S. K. Nayar, and J. J. Koenderink, "Reflectance and texture of real world surfaces," *ACM Trans. Graph.*, vol. 18, no. 1, pp. 1-34, 1999.
- [59] T. Ojala, et al., "Outex: new framework for empirical evaluation of texture analysis algorithm," in *Proc. Int. Conf. Pattern Recogn.*, pp. 701-706, Quebec City, Canada, 11-15 Aug. 2002.
- [60] A. R. Backes, D. Casanova, and O. M. Bruno, "Color texture analysis based on fractal descriptors," *Pattern Recognition*, vol. 45, no. 5, pp. 1984-1992, 2012.
- [61] S. Lazebnik, C. Schmid, and J. Ponce, "A sparse texture representation using local affine regions," *IEEE Trans Pattern Anal Mach Intell*, vol. 27, no. 8, pp. 1265-1278, Aug. 2005.
- [62] P. Brodatz, *Textures: A Photographic Album for Artists and Designers*, New York, NY, USA: Dover, 1966.
- [63] G. Smith, MeasTex Image Texture Database and Test Suite Centre for Sensor Signal and Information Processing, Univ Qld, 1998. <http://www.texturesynthesis.com/meastex/meastex.html>
- [64] G. J. Burghouts and J. M. Geusebroek, "Material-specific adaptation of color invariant features," *Pattern Recognition Letters*, vol. 30, no. 3, pp. 306-313, Feb. 2009.
- [65] J. Eastwood, L. Newton, R. Leach, and S. Piano, "Generation and categorization of surface texture data using a modified progressively growing adversarial network," *Precision Engineering*, vol. 74, pp. 1-11, Mar. 2022.
- [66] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image Process*, vol. 13, no. 4, pp. 600-612, Mar. 2004.
- [67] U. Lin, Z. Xu, G. Sharma, and T. N. Pappas, "Texture representation via analysis and synthesis with generative adversarial networks," *E-Prime-Advances in Electrical Engineering, Electronics and Energy*, vol. 6, Article ID: 100092, 2023.
- [68] S. E. Kim, H. Yoon, and J. Lee, "Fast and scalable earth texture synthesis using spatially assembled generative adversarial neural networks," *J. of Contaminant Hydrology*, vol. 243, no. C, Article ID: 103867, 2021.
- [21] I. J. Goodfellow, et al., "Generative adversarial nets," in *Proc. of the 27th Int. Conf. on Neural Information Processing Systems*, vol. 2, pp. 2672-2680, Lake Tahoe, NV, USA, 2-4 Dec. 2013.
- [22] M. Mirza and S. Osindero, *Conditional Generative Adversarial Nets*, arXiv preprint arXiv:1411.1784, 2014.
- [23] J.Y. Zhu, T. Park, P. Isola, A.A. Efros, *Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks*, arXiv preprint arXiv:1703.10593, 2017.
- [24] U. Bergmann, N. Jetchev, and R. Vollgraf, "Learning texture manifolds with the periodic spatial GAN," in *Proc. of the 34th Int. Conf. on Machine Learning*, vol. 70, pp. 469-477, Sydney, Australia, 6-11 Aug. 2017.
- [25] X. Li, G. Mariethoz, D. Lu, and N. Linde, "Patch-based iterative conditional geostatistical simulation using graph cuts," *Water Resource. Res.*, vol. 52, pp. 6297-6320, 2016.
- [26] A. Radford, L. Metz, and S. Chintala, Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks, arXiv preprint arXiv:1511.06434, 2015.
- [27] N. Jetchev, U. Bergmann, and R. Vollgraf, "Texture synthesis with spatial generative adversarial networks," arXiv preprint arXiv:1611.08207, 2016.
- [28] W. Fan, J. Fang, and G. Huang, "An improved image texture synthesis based on algorithm convolution neural network," *Physical Communication*, vol. 66, Article ID: 102395, 2024.
- [29] R. Paget and I. Longstaff, "Texture synthesis via a noncausal nonparametric multiscale Markov random field," *IEEE Trans. Image Process*, vol. 7, no. 6, pp. 925-931, Jun. 1998.
- [30] A. A. Efros and T. K. Leung, "Texture synthesis by non-parametric sampling," in *Proc. 7th Intl. Conf. Computer Vision*, vol. 2, pp. 1033-1038, Kerkyra, Greece, 20-27 Sept. 1999.
- [31] K. K. Singh and Y. J. Lee, "Hide-and-seek: forcing a network to be meticulous for weakly-supervised object and action localization," in *Proc. IEEE Int. Conf. on Computer Vision*, pp. 3544-3553, Venice, Italy, 22-29 Oct. 2017.
- [32] T. DeVries and G. W. Taylor, "Improved regularization of convolutional neural networks with cutout," arXiv preprint arXiv:1708.04552, 2017.
- [33] Z. Zhong, L. Zheng, G. Kang, S. Li, and Y. Yang, "Random erasing data augmentation," in *Proc. of the AAAI Conf. on Artificial Intelligence*, vol. 34, pp. 13001-13008, New York, NY, USA, 7-12 Feb. 2020.
- [34] P. Chen, S. Liu, H. Zhao, and J. Jia, "Gridmask data augmentation," arXiv preprint arXiv:2001.04086, 2020.
- [35] H. Inoue, Data Augmentation by Pairing Samples for Images Classification. arXiv preprint arXiv:1801.02929, 2018.
- [36] H. Zhang, M. Cisse, Y. N. Dauphin, and D. Lopez-Paz, "Mixup: beyond empirical risk minimization," arXiv preprint arXiv:1710.09412, 2017.
- [37] Y. Tokozume, Y. Ushiku, and T. Harada, "Between-class learning for image classification," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 5486-5494, Salt Lake City, UT, USA, 18-23 Jun. 2018.
- [38] S. Yun, D. Han, S. J. Oh, S. Chun, J. Choe, and Y. Yoo, "Cutmix: regularization strategy to train strong classifiers with localizable features," in *Proc. of the IEEE/CVF Int. Conf. on Computer Vision*, pp. 6023-6032, Seoul, South Korea, 27 Oct.-2 Nov. 2019.
- [39] T. DeVries and G.W. Taylor, *Improved Regularization of Convolutional Neural Networks with Cutout*, arXiv preprint arXiv:1708.04552, 2017.
- [40] A. Radford, L. Metz, and S. Chintala, Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks, arXiv preprint arXiv:1511.06434, 2015.
- [41] A. Madani, M. Moradi, A. Karargyris, and T. Syeda-Mahmood, "Chest x-ray generation and data augmentation for cardiovascular abnormality classification," *Medical Imaging 2018: Image Processing*, vol. 10574, pp. 105741M, 2018.
- [42] S. W. Huang, et al., "Auggan: cross domain adaptation with gan-based data augmentation," in *Proc. of the European Conf. on Computer Vision*, pp. 718-731, Munich, Germany, 8-14 Sept. 2018.
- [43] X. Zhu, Y. Liu, J. Li, T. Wan, and Z. Qin, "Emotion classification with data augmentation using generative adversarial networks," in *Pacific-Asia Conf. on Knowledge Discovery and Data Mining*, pp. 349-360, Melbourne, Australia, 3-6 Jun. 2018.
- [44] Z. Zheng, et al., "Generative adversarial network with multi-branch discriminator for imbalanced cross-species image-to-image translation," *Neural Netw.*, vol. 141, pp. 355-371, 2021.
- [45] A. Fawzi, H. Samulowitz, D. Turaga, and P. Frossard, "Adaptive data augmentation for image classification," in *Proc. IEEE Int. Conf. on Image Processing*, pp. 3688-3692, Phoenix, AZ, USA, 25-26 Sept. 2016.

محمدحسین شکور تحصیلات خود را در مقاطع کارشناسی و کارشناسی ارشد کامپیوتر به ترتیب در سال‌های ۱۳۷۷ و ۱۳۸۱ در دانشگاه‌های شیراز و اصفهان به پایان رساند. همچنین دکترای هوش مصنوعی خود را در سال ۱۳۹۵ از دانشگاه شیراز اخذ کرد. ایشان هم اکنون استادیار دانشگاه اراک در گروه کامپیوتر است و تخصص ایشان در زمینه پردازش تصویر، یادگیری عمیق و شناسایی الگو می باشد.

- [69] J. Portilla and E. P. Simoncelli, "A parametric texture model based on joint statistics of complex wavelet coefficients," *Int'l J. of Computer Vision*, vol. 40, no. 1, pp. 49-71, Oct. 2000.
- [70] U. Bergmann, N. Jetchev, and R. Vollgraf, "Learning texture manifolds with the periodic spatial GAN," in *Proc. of the 34th Int'l Conf. on Machine Learning Research*, vol. 70, pp. 469-477, Sydney, Australia 6-11 Aug. 2017.
- [71] Z. Guo, L. Zhang, and D. Zhang, "A completed modeling of local binary pattern operator for texture classification," *IEEE Trans. Image Process*, vol. 9, no. 16, pp. 1657-1663, Jun. 2010.