

Abstract

سازمان‌های بشردوستانه برای پاسخ به بلایایی مانند زلزله، آتش‌سوزی و درگیری‌های مسلحانه به داده‌های دقیق و به موقع در قالب ارزیابی خسارت نیاز دارند که نشان می‌دهد چه ساختمان‌ها و مراکز جمعیتی بیشترین آسیب را دیده‌اند. تحقیقات اخیر، یادگیری ماشینی را با سنجش از دور ترکیب می‌کند تا به طور خودکار چنین اطلاعاتی را از تصاویر ماهواره‌ای استخراج کند، کار دستی و زمان‌بر زدن را کاهش دهد. یکی از موانع اصلی استفاده از روش‌های یادگیری ماشین در سناریوهای واکنش به بلایای واقعی، دشواری به دست آوردن مقدار کافی از داده‌های برچسب‌گذاری شده برای آموزش مدلی برای یک فاجعه آشکار است. این مقاله یک کاربرد جدید از یادگیری نیمه نظارت شده (SSL) را برای آموزش مدل‌هایی برای ارزیابی آسیب با حداقل مقدار داده برچسب‌دار و مقدار زیادی از داده‌های بدون برچسب نشان می‌دهد. ما عملکرد روش‌های پیشرفته SSL، از جمله MixMatch [28] و FixMatch [28] را با یک خط پایه نظارت شده برای زلزله 2010 هائیتی، آتش‌سوزی سانتا روزا 2017 [15] و درگیری مسلحانه 2016 در سوریه [32] مقایسه می‌کنیم. ما نشان می‌دهیم که چگونه مدل‌های آموزش‌دیده شده با روش‌های SSL می‌توانند علیرغم استفاده از تنها بخشی از داده‌های برچسب‌گذاری شده به عملکرد کاملاً نظارت شده برسند و مناطقی را برای بهبود بیشتر شناسایی می‌کنیم.

Introduction

هنگامی که یک بحران انسانی مانند یک بلای طبیعی رخ می‌دهد، پاسخ دهندگان به بحران باید مکان جمعیت‌های آسیب‌دیده را برای تسهیل تلاش‌های امدادی بدانند. مکان‌ها و تراکم ساختمان‌های آسیب‌دیده به عنوان یک پروکسی مفید برای تخمین این اطلاعات عمل می‌کند [6]. یک رویکرد برای شناسایی آنها سنجش از دور است: تحلیلگران متخصص تصاویر ماهواره‌ای قبل و بعد از فاجعه را از منطقه آسیب‌دیده مقایسه می‌کنند و مکان ساختمان‌های آسیب‌دیده را علامت‌گذاری می‌کنند. با این حال، این زمان‌بر است (کمتر از 100 ساختمان در ساعت برای هر نفر ارزیابی می‌شود)، مقیاس‌پذیری چالش برانگیز و مستعد خطا است [29، 19]. یادگیری ماشینی (ML) به عنوان یک ابزار کارآمد برای خودکارسازی فرآیند ارزیابی آسیب استفاده شده است [4، 18، 9، 15، 32، 16، 30]. مدل‌ها برای تمایز بین تصاویر ساختمان‌های آسیب‌دیده و آسیب‌دیده با استفاده از تصاویر برچسب‌گذاری شده توسط کارشناسان از بلایای گذشته آموزش داده می‌شوند. مدل‌های آموزش‌دیده می‌توانند کل شهرها را در چند دقیقه زمانی که در مراکز داده مدرن مستقر شوند، تجزیه و تحلیل کنند. ساخت مدل‌های دقیق برای بلایای جدید همچنان چالش برانگیز است. ML به طور سنتی بر مجموعه داده‌هایی با تنوع و پوشش کافی تکیه می‌کند، به طوری که مدل‌ها در زمان استنتاج به نمونه‌های نادیده تعمیم می‌یابند. این در حوزه ما امکان‌پذیر نیست زیرا تعداد محدودی از بلایای گذشته وجود دارد و ظاهر هر فاجعه یا منطقه جغرافیایی به طور گسترده‌ای در چیدمان ساختمان‌ها متفاوت است.

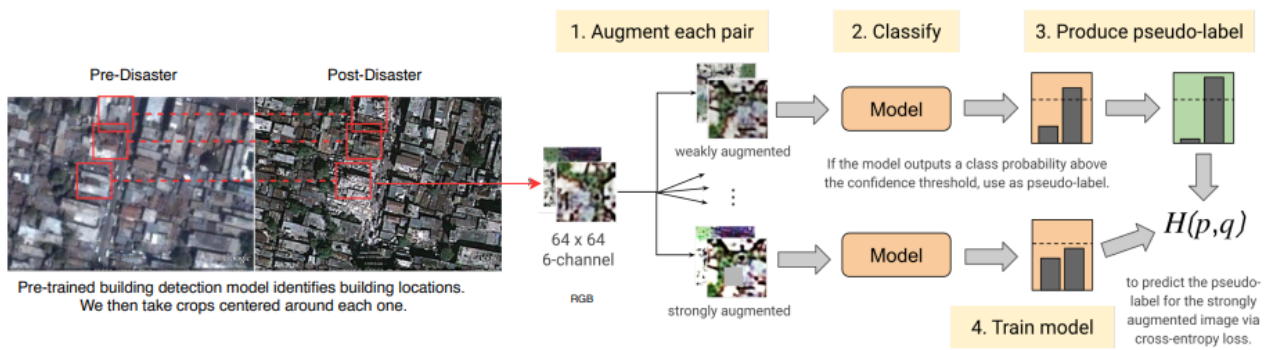
هنگامی که یک بحران انسانی مانند یک بلای طبیعی رخ می‌دهد، پاسخ دهندگان به بحران باید مکان جمعیت‌های آسیب‌دیده را برای تسهیل تلاش‌های امدادی بدانند. مکان‌ها و تراکم ساختمان‌های آسیب‌دیده به عنوان یک پروکسی مفید برای تخمین این اطلاعات عمل می‌کند [6]. یک رویکرد برای شناسایی آنها سنجش از دور است: تحلیلگران متخصص تصاویر ماهواره‌ای قبل و بعد از فاجعه را از منطقه آسیب‌دیده مقایسه می‌کنند و مکان ساختمان‌های آسیب‌دیده را علامت‌گذاری می‌کنند. با این حال، این زمان‌بر است (کمتر از 100 ساختمان در ساعت برای هر نفر ارزیابی می‌شود)، مقیاس‌پذیری چالش برانگیز و مستعد خطا است [29، 19]. یادگیری ماشینی (ML) به عنوان یک ابزار کارآمد برای خودکارسازی فرآیند ارزیابی آسیب استفاده شده است [4، 18، 9، 15، 32، 16، 30]. مدل‌ها برای تمایز بین تصاویر ساختمان‌های آسیب‌دیده و آسیب‌دیده با استفاده از تصاویر برچسب‌گذاری شده توسط کارشناسان از بلایای گذشته آموزش داده می‌شوند. مدل‌های آموزش‌دیده می‌توانند کل شهرها را در چند دقیقه زمانی که در مراکز داده مدرن مستقر شوند، تجزیه و تحلیل کنند. ساخت مدل‌های دقیق برای بلایای جدید همچنان چالش برانگیز است. ML به طور سنتی بر مجموعه داده‌هایی با تنوع و پوشش کافی تکیه می‌کند، به طوری که مدل‌ها در زمان استنتاج به نمونه‌های نادیده تعمیم می‌یابند. این در حوزه ما امکان‌پذیر نیست زیرا تعداد محدودی از بلایای گذشته وجود دارد و ظاهر هر فاجعه یا منطقه جغرافیایی به طور گسترده‌ای در طرح ساختمان‌ها، مصالح ساختمانی، پوشش گیاهی، ظاهر آسیب و غیره متفاوت است [24] (شکل 3 را در ضمیمه ببینید). علاوه بر این، حتی اگر مدل در همان نوع فاجعه و مکان از قبل

آموزش داده شده باشد، به دلیل تغییر مناظر، تغییرات فصلی، پوشش ابر و سایر عواملی که باعث می‌شود داده‌های استخراج به طور سیستماتیک با داده‌های آموزشی متفاوت باشد، نویز ذاتی در تصاویر ماهواره‌ای وجود دارد. بنابراین، مدل‌هایی که فقط در مورد بلایای گذشته آموزش دیده‌اند، احتمالاً در بلایای جدید ضعیف عمل خواهند کرد. ما می‌توانیم با آموزش مدل‌هایی بر روی داده‌های یک فاجعه جدید پس از وقوع آن، از مشکل تعمیم جلوگیری کنیم. مدل‌هایی که به این روش آموزش داده می‌شوند باید فقط از تعداد کمی از نمونه‌های آموزشی برچسب‌گذاری شده استفاده کنند، زیرا برچسب‌گذاری دستی متخصص زمان‌بر است. آغاز یک فاجعه جدید، داده‌های برچسب‌گذاری شده محدود است، اما مقدار زیادی از تصاویر ماهواره‌ای بدون برچسب را می‌توان به طور خودکار از منطقه استخراج کرد. پیشرفت‌های اخیر در تکنیک‌های یادگیری نیمه‌نظارت‌شده (SSL) نشان می‌دهد که الگوریتم‌هایی که داده‌های آموزشی برچسب‌گذاری شده و بدون برچسب را ترکیب می‌کنند، می‌توانند عملکردی قابل مقایسه با الگوریتم‌های کاملاً نظارت‌شده آموزش دیده‌شده بر اساس داده‌های برچسب‌گذاری شده‌تر را به دست آورند [2، 28، 1]. با استفاده از تکنیک‌های SSL، می‌توانیم مدل‌های ارزیابی دقیق آسیب را برای بلایای جدید بدون صرف زمان زیادی برای جمع‌آوری برچسب‌های دستی آموزش دهیم. در این مقاله، ما از دو تکنیک [2] MixMatch، SSL و [28] FixMatch برای آموزش مدل‌های تشخیص آسیب ساختمان با استفاده از مقدار محدودی از داده‌های برچسب‌گذاری شده استفاده می‌کنیم. ما سه فاجعه مختلف را ارزیابی می‌کنیم که تعداد نمونه‌های برچسب‌گذاری شده را متفاوت می‌دانیم. ما نشان می‌دهیم که مدل‌ها می‌توانند تنها با استفاده از کسری از داده‌های برچسب‌گذاری شده به سطح نتایج کاملاً نظارت شده نزدیک شوند.

Related Work

یادگیری ماشین در ارزیابی ساختمان آسیب مطالعات گذشته از تنظیمات کاملاً نظارت شده با موفقیت از روش‌های یادگیری ماشین برای تشخیص آسیب ساختمان از تصاویر ماهواره‌ای استفاده کرده است. مجموعه داده عمومی [15، 13] xBD در کنار چالش [14] xView2 منتشر شد، که تصاویر ماهواره‌ای در مقیاس بزرگ، چند ضلعی‌های ساختمانی و برچسب‌های ترتیبی ارائه می‌دهد که سطح آسیب را در 19 فاجعه با وظیفه طبقه‌بندی هر پیکسل نشان می‌دهد. رویکرد مقام اول [10] دو مرحله داشت، در ابتدا یک مدل بومی‌سازی را با تصاویر قبل از فاجعه آموزش داد و سپس از وزن‌ها برای راه‌اندازی یک شبکه عصبی سیامی برای طبقه‌بندی ساختمان استفاده کرد که وزن‌ها را بین تصاویر قبل از فاجعه و پس از فاجعه تقسیم می‌کند. گوپتا و همکاران [16] یک رویکرد انتها به انتها را پیشنهاد کرد، ابتدا ویژگی‌های تصویر در مقیاس چندگانه را استخراج کرد، آنها را در یک سر تقسیم بندی تغذیه کرد تا ساختمان‌ها را به طور مستقل بر روی تصاویر قبل و بعد از فاجعه پیش‌بینی کند، و در نهایت هر پیکسل را طبقه‌بندی کرد. وبر و همکاران [30] یک شبکه واحد را آموزش داد و کار را به عنوان بخش بندی معنایی مدلسازی کرد. روش‌های بالا داده‌های آموزشی و اعتبارسنجی را از توزیع یکسان تولید می‌کنند، بنابراین مشکل اجرای استنتاج برای یک فاجعه جدید با حداقل داده را برطرف نمی‌کنند. خو و همکاران [32] مدل‌هایی را برای زلزله‌های هائیتی، مکزیک و اندونزی توسعه دادند و آزمایش‌های تعمیم بین منطقه‌ای را انجام دادند، که نشان می‌داد چگونه مدل‌های از قبل آموزش دیده‌شده در مورد بلایای گذشته در منطقه جدید بدون داده‌های برچسب‌گذاری شده یا حداقل داده‌های برچسب‌گذاری شده خوب عمل نمی‌کنند. ما از این روش به عنوان پایه در آزمایشات خود استفاده می‌کنیم. والتجین و همکاران [29] آزمایش‌هایی را روی 13 بلایای مختلف انجام داد که از نظر نوع خطر، منطقه جغرافیایی و پارامترهای ماهواره‌ای متفاوت بودند و متوجه شدند که عملکرد به طور قابل‌توجهی در بین بلایای آزمایشی متفاوت است، خواه داده‌های فاجعه آزمایشی در آموزش گنجانده شده باشد یا نباشد. رویکردهای یادگیری نیمه نظارتی یادگیری نیمه نظارتی (SSL) رویکردهایی را برای کاهش نیاز به مقادیر زیادی از داده‌های برچسب دار با استفاده از داده‌های بدون برچسب ارائه می‌دهد. دسته‌ای از روش‌های SSL برای شبکه‌های عمیق وجود دارد که برچسب‌گذاری کاذب (یا خودآموزی) را انجام می‌دهند [23، 25، 31، 27]، برچسب‌های مصنوعی را از داده‌های بدون برچسب تولید می‌کنند و شامل حداقل نیروی انسانی می‌شوند [34، 2، 1، 28]. این تکنیک‌ها بر منظم‌سازی سازگاری [26، 22] تکیه می‌کنند، که مدل را تشویق می‌کند تا پیش‌بینی‌های یک توزیع مشابه را در میان آشفته‌گی‌های یک ورودی داده شده ارائه دهد. [2] MixMatch انواع دیگری از منظم‌سازی را اضافه می‌کند، با استفاده از [34] MixUp برای تشویق رفتار محدب «بین» نمونه‌ها با تولید ترکیب‌های وزن‌دار از نمونه‌های برچسب‌دار و بدون برچسب. [28] FixMatch یک رویکرد ساده‌تر ارائه کرد که به عملکرد پیشرفته‌ای در معیارهای رایج SSL، مانند [20] CIFAR و [3] STL دست یافت. روش‌ها در نحوه استفاده از برچسب‌های شبه برای محاسبه ضرر متفاوت است. برای یک تصویر بدون برچسب معین، MixMatch یک برچسب حدس را بر اساس نسخه‌های ضعیف افزوده شده آن ایجاد می‌کند و ضرر را بر اساس میزان پیش‌بینی مدل آن برچسب را محاسبه می‌کند. FixMatch همچنین یک برچسب شبه از تقویت‌های ضعیف ایجاد می‌کند، اما ضرر را بر این اساس محاسبه

می‌کند که آیا مدل قادر به پیش‌بینی برچسب در نسخه‌های تقویت‌شده قوی است یا خیر. مفروضات محرک روش های شبه برچسب گذاری SSL باید برای تصاویر ماهواره ای صادق باشد و راهی برای استفاده از داده های بدون برچسب ارائه دهد.



شکل 1: خط لوله که نشان می دهد چگونه تصاویر قبل و بعد از فاجعه در یک ورودی 6 کانالی انباشته شده و سپس برای تولید برچسب های شبه برای آموزش افزوده می شوند. این نمودار شامل تقویت قوی برای نشان دادن FixMatch است، اما MixMatch فقط از تصاویر تقویت شده ضعیف استفاده می کند.

Data

ما رویکرد خود را در مورد سه فاجعه ارزیابی کردیم. علاوه بر آتش سوزی سانتا روزا [15]، ما مجموعه داده های خود را برای زلزله 2010 هائیتی و عکس فوری حلب در سال 2016 مانند [32] تولید کردیم. ابتدا، ما تصاویری را از قبل و بعد از فاجعه برای هر منطقه، عمدتاً از ماهواره های WorldView 2 و DigitalGlobe 3 بدست آوردیم. برای هائیتی، تصاویر صریح از پرواز توسط اداره ملی اقیانوسی و جوی ارائه شده است. ما همه تصاویر را با وضوح 0.3 متر برای ثبات نمونه برداری کردیم. در مرحله بعد، از ارزیابی های آسیب ساختمان ارائه شده توسط UNOSAT، برنامه کاربردی ماهواره ای عملیاتی مؤسسه آموزش و تحقیقات سازمان ملل متحد (UNITAR)، که در وبسایت تبادل داده های بشردوستانه موجود است، برچسب های واقعی مثبت به دست آوردیم [17]. ارزیابی های UNOSAT از یک مقیاس 5 سطحی برای اندازه گیری آسیب استفاده می کنند، اما برچسب ها در مجموعه داده های مختلف پر سر و صدا و ناسازگار بودند. بنابراین، ما "Severe" Damage و "Destroyed" را در یک کلاس "Damaged" گروه بندی کردیم و مشکل خود را به عنوان یک مشکل طبقه بندی باینری برای شناسایی ساختمان های "Damaged" و "Undamaged" فرموله کردیم. برای به دست آوردن نمونه های آسیب ندیده، از یک مدل تشخیص ساختمان از پیش آموزش دیده [32] برای شناسایی ساختمان ها استفاده کردیم و ساختمان هایی که در ارزیابی های UNOSAT به عنوان آسیب دیده مشخص شده بودند را فیلتر کردیم. در نهایت، برای ایجاد نمونه های آموزشی، از محصولاتی که در اطراف هر ساختمان قرار داشتند نمونه برداری کردیم. سپس تصاویر قبل و بعد از فاجعه را تراز کردیم و از Google Earth Engine [12] برای پیوستن مکانی به برچسب ها و تصاویر برش خورده استفاده کردیم. هر نمونه در مجموعه داده ما حاوی یک تصویر 6 کانالی با ابعاد 64 x 64 و یک برچسب طبقه بندی است (0 برای آسیب نخورده، 1 برای آسیب دیده). 50742 نمونه هائیتی (44 درصد مثبت)، 12897 نمونه سانتا روزا (27 درصد مثبت) و 10452 نمونه حلب (44 درصد مثبت) وجود دارد. در یک فاجعه جدید، این حجم از داده های برچسب گذاری شده در دسترس نخواهد بود. برای شبیه سازی این، یک نمونه تصادفی از نمونه ها با برچسب و بقیه بدون برچسب در نظر گرفته شد (بخش 5).

Approach

ما وظیفه خود را به عنوان طبقه بندی باینری فرموله می کنیم، که در آن دو کلاس بدون آسیب (0) یا آسیب دیده (1) هستند. ما دسته ای از نمونه های برچسب گذاری شده B را به صورت $\{(x_b, p_b) \mid b \in \{0, 1\}\}$ تعریف می کنیم، که در آن x_b شامل تصاویر نمونه 6 کانالی است و p_b برچسب های تک داغ هستند. ما همچنین دسته ای از نمونه های بدون برچسب μB را به عنوان $U = \{u_b \mid u_b \in \{0, 1\}\}$ تعریف می کنیم، که μ یک فرارامتر است که اندازه های نسبی X و U را تعیین می کند. ما مدل های طبقه بندی را با استفاده از MixMatch و FixMatch آموزش دادیم، که از شبه برچسب گذاری برای ایجاد برچسب های مصنوعی برای U استفاده می کنند. سپس، یک مثال آموزش داده شده به شبکه ای CNN تولید می شود. از هر دو X

و U ، با استفاده از یک تابع ضرر که دارای یک عبارت ضرر برچسب دار و بدون برچسب است. ما تفاوت‌ها را در توابع شبه برچسب‌گذاری و از دست دادن در زیر مقایسه می‌کنیم. تولید برچسب‌های شبه ما ابتدا مرحله حدس زدن برچسب MixMatch [2] را بررسی می‌کنیم. MixMatch در ابتدا تقویت ضعیف را اعمال می‌کند، که شامل چرخش‌ها، چرخش‌ها و جابجایی‌های تصادفی می‌شود، هم به X و هم U . اجازه دهید $\alpha(\cdot)$ نشان‌دهنده تقویت ضعیف باشد، به‌طوری‌که، برای یک دسته ub بدون برچسب معین، $\alpha(ub)$ حاوی افزایش‌های K برای هر مثال باشد. ما میانگین توزیع‌های کلاس پیش‌بینی‌شده مدل را در تمام تقویت‌ها محاسبه می‌کنیم که $p_{\text{model}}(y|\alpha(ub, k))$ است. سپس، روش تیز کردن را اعمال می‌کند، که آنتروپی توزیع برچسب را همانطور که در [11] معرفی شد کاهش می‌دهد. $qb = \text{Sharpen}(\bar{qb})$ به عنوان هدف برای پیش‌بینی مدل در افزایش ub عمل می‌کند.

در نهایت، MixMatch نسخه‌های افزوده‌شده داده‌های برچسب‌گذاری شده $\alpha(X)$ و $\alpha(ub)$ را می‌گیرد و هم مخلوط کردن را اعمال می‌کند. ما به مجموعه داده برچسب‌گذاری‌شده نهایی با افزایش‌ها و MixUp به عنوان X_0 و مجموعه داده بدون برچسب نهایی به عنوان U_0 اشاره می‌کنیم. اکنون، ما مرحله شبه برچسب‌گذاری [28] FixMatch را بررسی می‌کنیم که روش نسبتاً ساده‌تری را ارائه می‌دهد اما در مجموعه داده‌های معیار عملکرد بهتری داشت. مانند MixMatch، FixMatch اهداف را از نسخه‌های ضعیف تقویت شده یک مثال بدون برچسب ایجاد می‌کند. با این حال، بدون تیز کردن توزیع خروجی، از $\arg \max(qb)$ به عنوان شبه برچسب استفاده می‌کند. FixMatch نه تنها از تقویت ضعیف بلکه از تقویت قوی نیز استفاده می‌کند که با $A(\cdot)$ نشان داده می‌شود. در این کار، ما از FixMatch با [1] RandAugment (RA) استفاده می‌کنیم که به طور تصادفی یک سری تبدیل اعمال می‌کند. انواع دگرگونی‌ها شامل تغییر روشنایی، کنتراست یا سطح اشباع، خورشیدی کردن، پوستر کردن، اعمال CutOut و موارد دیگر بود [28]. FixMatch با تشویق مدل به پیش‌بینی برچسب شبه برای نسخه‌های قویاً تقویت‌شده ub یا $A(ub)$ از تصاویر ضعیف و قوی استفاده می‌کند. ما پیامدهای این تفاوت را با مقایسه توابع ضرر در زیر نشان می‌دهیم.

تابع ضرر برای هر دو روش، تابع ضرر دارای یک عبارت زیان برچسب دار یا نظارت شده، L_s ، و یک عبارت ضرر بدون برچسب، L_u است. آنها در یک هدف آموزشی به صورت $L = L_s + \lambda L_u$ ترکیب می‌شوند، که در آن λ یک فرایارامتر است که وزن نسبی از دست دادن بدون برچسب را نشان می‌دهد. از دست دادن برچسب برای هر دو روش مشابه است. در MixMatch، برای هر دسته $x_0 \in X_0$ و $b \in B$ ، تلفات برچسب‌گذاری شده $H(p_b, p_{\text{model}}(y|x_0))$ است، که در آن $H(p, q)$ به عنوان آنتروپی متقابل بین دو توزیع احتمال تعریف می‌شود. در FixMatch، جایی که MixUp اعمال نمی‌شود، $L_s = 1/B \sum_b H(p_b, p_{\text{model}}(y|\alpha(x_b)))$. مدت ضرر بدون برچسب به طور قابل توجهی متفاوت است. برای MixMatch، مجذور ضرر L_2 در پیش‌بینی‌ها و برچسب‌های حدس‌شده است که با معادله نشان داده شده است. (1) برای هر دسته $u \in U_0$.

$$\mathcal{L}_u = \frac{1}{\mu B} \sum_{b=1}^B \|q_b - p_{\text{model}}(y|u'_b)\|_2^2 \quad (1)$$

Alternatively, FixMatch enforces cross-entropy loss against the strongly augmented unlabeled examples given the weakly augmented ones as shown in eq. (2).

$$\mathcal{L}_u = \frac{1}{\mu B} \sum_{b=1}^B \mathbb{1}(\max(q_b) \geq \tau) H(\arg \max(q_b), p_{\text{model}}(y|\mathcal{A}(u_b))) \quad (2)$$

که در آن τ یک فرایارامتر اسکالر است که آستانه‌ای را نشان می‌دهد که بالای آن یک برچسب شبه را حفظ می‌کنیم. اگرچه رویکرد Fixmatch ساده‌تر است و از نظر تجربی بهتر عمل می‌کند، ما آزمایش‌هایی را با هر دو روش انجام دادیم زیرا تغییر دامنه به تصاویر ماهواره‌ای است، جایی که تقویت‌ها ممکن است مفاهیم معنایی متفاوتی داشته باشند. به عنوان مثال، اعمال یک افزایش ضعیف یک جابجایی تصادفی ممکن است باعث شود که مدل یک ساختمان را به عنوان آسیب دیده پیش‌بینی کند زیرا ساختمان‌ها می‌توانند در زلزله جابجا شوند. روش دیگر، به دلیل ماهیت پرسر و صدا تصاویر ماهواره‌ای، تقویت‌های قوی می‌تواند یادگیری مدل را چالش‌برانگیزتر کند. بنابراین، ما هر دو MixMatch و FixMatch را در آزمایشات خود امتحان کردیم.

Experiments & Results

ما کارایی روش‌های SSL را بر طبق‌بندی ساختمان‌ها در سه منطقه مختلف آزمایش می‌کنیم: سانتا روزا [15]، هائیتی و حلب [32]، با استفاده از چهار روش مختلف: مدل برج دوقلو که در کار گذشته Xu و همکارانش بهترین عملکرد را داشت. [32] به عنوان پایه، یادگیری کاملاً نظارت شده تنها با داده‌های آموزشی برچسب‌گذاری شده، و یادگیری نیمه نظارتی با MixMatch [2] و [28] FixMatch. تنظیم. ما مجموعه داده‌ها را به ترتیب برای قطار و آزمایش به 90% و 10% تقسیم کردیم. برای آزمایش‌های SSL، تعداد خاصی از مثال‌ها (مانند 10، 50، 100، 500) را به‌طور تصادفی از هر کلاس به‌عنوان مجموعه آموزشی برچسب‌گذاری‌شده نمونه‌برداری می‌کنیم و بقیه را به‌عنوان مجموعه بدون برچسب در نظر می‌گیریم. ما آزمایش‌ها را 5 بار با تقسیم‌های مختلف مجموعه آموزشی برچسب‌گذاری شده انجام می‌دهیم. ما یک نوع شبکه باقی مانده گسترده (WRN) [33] را آموزش می‌دهیم که شامل چهار بلوک باقیمانده با 32، 64، 128، و 256 فیلتر است. ما از نزدیک استراتژی‌های آموزشی را در [28] دنبال می‌کنیم، با استفاده از تکانه SGD (0.9) با کاهش نرخ یادگیری کسینوس. ما فرآیندهای دیگر، مانند نرخ یادگیری یا کاهش وزن، را برای هر روش تنظیم می‌کنیم، اما آنها در بین مجموعه داده‌ها به اشتراک گذاشته می‌شوند. ما جزئیات تجربی را در مواد تکمیلی ارائه می‌دهیم. نتایج در جدول 1، دقت را به‌طور میانگین در 5 اجرا گزارش می‌کنیم که هر کدام بر روی تقسیم‌بندی داده‌های برچسب‌گذاری شده متفاوت آموزش داده شده‌اند. ما همچنین یک مدل کاملاً نظارت شده را روی همه داده‌های برچسب‌گذاری شده (به نام "۹۰٪ نظارت شده") به عنوان کران بالای عملکرد آموزش می‌دهیم، اگرچه این در عمل در دسترس نخواهد بود.

Dataset	# Labeled Data	Twin Tower	Fully Supervised	MixMatch	FixMatch
Haiti	10	0.53 ± 0.03	0.58 ± 0.02	0.60 ± 0.11	0.56 ± 0.05
	50	0.56 ± 0.02	0.64 ± 0.01	0.72 ± 0.03	0.61 ± 0.03
	100	0.56 ± 0.05	0.69 ± 0.02	0.75 ± 0.04	0.75 ± 0.10
	500	0.71 ± 0.01	0.75 ± 0.01	0.82 ± 0.01	0.87 ± 0.01
	90% supervised (45,667 data)	0.90	-	-	-
Santa Rosa	10	0.54 ± 0.03	0.74 ± 0.08	0.70 ± 0.06	0.92 ± 0.07
	50	0.54 ± 0.03	0.91 ± 0.03	0.85 ± 0.08	0.96 ± 0.02
	100	0.58 ± 0.04	0.92 ± 0.01	0.88 ± 0.04	0.97 ± 0.01
	500	0.69 ± 0.03	0.96 ± 0.01	0.96 ± 0.01	0.98 ± 0.00
	90% supervised (11,067 data)	0.99	-	-	-
Aleppo	10	0.52 ± 0.04	0.55 ± 0.05	0.65 ± 0.08	0.72 ± 0.15
	50	0.56 ± 0.02	0.65 ± 0.06	0.73 ± 0.06	0.88 ± 0.01
	100	0.57 ± 0.04	0.71 ± 0.06	0.78 ± 0.04	0.89 ± 0.01
	500	0.67 ± 0.04	0.82 ± 0.01	0.85 ± 0.02	0.90 ± 0.01
	90% supervised (9,406 data)	0.88	-	-	-

Table 1: Classification accuracy of the Twin Tower [32] baseline, fully-supervised, MixMatch, and FixMatch models with varying amounts of labeled training data, averaged over 5 runs (each with a different labeled data split). To provide an upper bound, one model is trained using all training data.

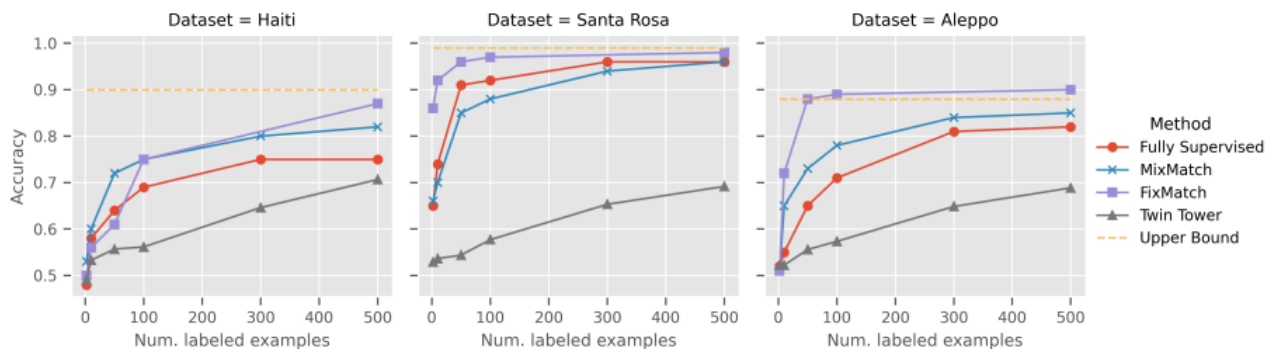


Figure 2: Performance of fully and semi-supervised models by number of labeled training examples.

این SSL دقت را در هر سه مجموعه داده فاجعه بهبود می بخشد. MixMatch بدون توجه به تعداد داده‌های آموزشی برچسب‌گذاری‌شده، پیشرفت‌های ثابتی را نسبت به برج‌های دوقلو و مدل‌های کاملاً تحت نظارت در هائیتی و حلب نشان می‌دهد. FixMatch در هر سه منطقه بهبودهای قابل توجهی ایجاد می‌کند، اگرچه برای پیشی گرفتن از عملکرد MixMatch در هائیتی به 100 یا بیشتر داده برچسب زده شده نیاز دارد. در حلب، FixMatch با 500 داده برچسب‌گذاری شده می‌تواند از مدل کاملاً نظارت شده آموزش‌دیده بر روی تمام داده‌های برچسب‌گذاری شده بهتر عمل کند، احتمالاً به این دلیل که تقویت‌های قوی سطحی از واریانس را نه در مجموعه داده‌ها به تنهایی ثبت می‌کنند. ما سیاست‌های تقویت را در ضمیمه C بررسی می‌کنیم. به طور کلی، نتایج ما عمومیت روش‌های SSL مدرن مبتنی بر تقویت داده‌ها را نشان می‌دهد، مانند MixUp [34]، CTAugment [1]، RandAugment [5]، علی‌رغم آمار تصویر متضاد تصاویر ماهواره‌ای در مقایسه با معیارهای تشخیص بصری استاندارد، مانند CIFAR-10، مانند CIFAR-10.

Conclusion

در این مقاله، ما یک کاربرد جدید از یادگیری نیمه نظارتی را برای تشخیص خودکار ساختمان‌های آسیب‌دیده در تصاویر ماهواره‌ای با داده‌های برچسب‌دار محدود معرفی کردیم. ما با دو تکنیک اخیر، MixMatch و FixMatch آزمایش کردیم و نشان دادیم که چگونه می‌توانند با استفاده از داده‌های بدون برچسب، به عملکرد قوی 100 نمونه برچسب‌دار یا کمتر دست پیدا کنند. آنها به طور مداوم از مدل‌های کاملاً نظارت شده بهتر عمل کردند و حتی به عملکردی نزدیک به تنظیمات کاملاً نظارت شده و بدون محدودیت داده دست یافتند. نتایج به‌طور تجربی نشان داد که چگونه رویکردهای SSL می‌تواند برای آموزش مدل‌ها مفید باشد، زمانی که یک فاجعه جدید در یک منطقه نامرئی در حال آشکار شدن است. برای کارهای آینده، ما قصد داریم نحوه ترکیب موثر داده‌های بلایای گذشته را بررسی کنیم. ممکن است دگرگونی‌های مستقل از منطقه ناشی از فاجعه‌ای باشد که مدل‌ها به اندازه کافی آن را ثبت نمی‌کنند یا انواع مختلفی از افزایش‌ها و تلفات که در برابر نویز ذاتی تصاویر ماهواره‌ای قوی‌تر هستند.

Acknowledgments and Disclosure of Funding

این کار با همکاری Google Research و شتاب دهنده نوآوری برنامه جهانی غذای سازمان ملل متحد (WFP) است. شتاب دهنده نوآوری WFP راه حل‌های بالقوه بالا برای گرسنگی را در سراسر جهان شناسایی، پشتیبانی و مقیاس می‌کند. ما از مبتکران برنامه جهانی غذا و استارت‌آپ‌ها و شرکت‌های خارجی از طریق حمایت مالی، دسترسی به شبکه‌ای از کارشناسان و دسترسی به حوزه جهانی حمایت می‌کنیم. ما معتقدیم راه رو به جلو در مبارزه با گرسنگی لزوماً ساختن برنامه‌های بزرگ نیست، بلکه شناسایی و آزمایش راحل‌ها به روشی چابک است. شتاب دهنده نوآوری فضایی است که در آن جهان می‌تواند بفهمد چه چیزی در مقابله با گرسنگی مفید است و چه چیزی مفید نیست - مکانی که می‌توانیم جسور باشیم، شکست بخوریم و همچنین موفق شویم.