

Damage Assessment in Military Operations using Deep Learning and Images Processing

Mohammad Ali Hamidi¹, Hamid Bigdeli^{2*}

1- Institute for the Study of War, Army Command and Staff University, Tehran, I.R.Iran.

2*- Institute for the Study of War, Army Command and Staff University, Tehran, I.R.Iran.

¹ Mohammad69h94@gmail.com, ^{2*} Hamidbigdeli92@gmail.com

Corresponding author's address: Hamid Bigdeli, Faculty Member, Army Command and Staff University, Tehran, Iran.

Abstract- Object detection plays an important role in the analysis of images of battle scenes, especially damage assessment. In this article, machine vision techniques along with deep learning and digital image processing are used to achieve a powerful method with maximum speed and accuracy to detect objects on the battlefield and also estimate the damage. To object detection, the features and parameters of images are extracted by convolutional neural networks and used in neural network learning. The structural similarity criteria, mean square error, and threshold method were used to assess the damage and to measure the similarity and changes in the images which received before and after military operations. Finally, for validating the method, samples of the battle scenes images have been investigated, and object detection and damage assessment has been executed on them.

Keywords- Image assessment, Object detection, Image processing, Deep learning.

ارزیابی خسارت در عملیات نظامی با استفاده از یادگیری عمیق و پردازش تصاویر

محمدعلی حمیدی^۱، حمید بیگدلی^{۲*}

۱- پژوهشکده عالی جنگ، دانشگاه فرماندهی و ستاد آجا، تهران، ایران.

۲* - پژوهشکده عالی جنگ، دانشگاه فرماندهی و ستاد آجا، تهران، ایران.

¹Mohammad69h94@gmail.com, ^{2*}Hamidbigdeli92@gmail.com

تاریخ دریافت: تاریخ بازنگری: تاریخ پذیرش:

* نشانی نویسنده مسئول: حمید بیگدلی، تهران، دانشگاه فرماندهی و ستاد آجا، پژوهشکده عالی جنگ، کد پستی: ۴۷۱۴۸-۷۱۱۶۷

چکیده- تشخیص اشیاء نقش مهمی در تحلیل تصاویر صحنه‌های نبرد به ویژه ارزیابی خسارت دارد. در این مقاله از فنون بینایی ماشین و در کنار آن یادگیری عمیق و پردازش تصاویر دیجیتال برای دستیابی به روشی قدرتمند با حداکثر سرعت و دقت جهت تشخیص اشیاء مورد نظر در صحنه نبرد و همچنین تخمین خسارات وارد شده استفاده شده است. برای انجام عمل تشخیص اشیاء، ویژگی‌ها و پارامترهای تصاویر توسط شبکه‌های عصبی کانولوشن استخراج شده و در امر یادگیری شبکه عصبی به کار گرفته شده است. برای ارزیابی خسارت از معیارهای شباهت ساختاری، خطای میانگین مربعات و روش آستانه‌گذاری برای سنجش میزان شباهت و تغییرات تصاویر دریافتی قبل و بعد از خسارات وارد شده، استفاده شده است. در نهایت، نمونه‌هایی از تصاویر مکان‌های خسارت دیده برای اعتبارسنجی روش پیشنهادی مورد بررسی و آزمایش قرار گرفته است.

واژه‌های کلیدی: ارزیابی خسارت، تشخیص اشیاء، پردازش تصویر، بینایی ماشین، یادگیری عمیق

۱- مقدمه

اگرچه مسائل تشخیص اشیاء^۱ و ارزیابی خسارت در عملیات‌های نظامی موضوعاتی قدیمی هستند و تلاش‌های زیادی برای پیشرفت آن‌ها صورت گرفته؛ اما همچنان مسائلی پیچیده هستند و زمینه‌ای فعال برای بسیاری از پژوهشگران می‌باشد. با وجود پیشرفت‌های زیاد هنوز روش‌های موجود دقت و سرعت کافی جهت استفاده در بسیاری از کاربردهای پیچیده واقعی را ندارند. یکی از رویکردهای مهم برای بررسی این مسأله استفاده از روش‌های هوش مصنوعی^۲ است. هوش مصنوعی، شاخه‌ای از علوم کامپیوتر است که دربرگیرنده تمام فرآیندهایی است که به شبیه‌سازی هوش انسان بر روی کامپیوترها و یا دستگاه‌هایی که توسط کامپیوتر کنترل می‌شوند، می‌پردازد. امروزه پژوهش‌های هوش مصنوعی به‌صورت عمده، بر جنبه‌های یادگیری^۳، استدلال^۴، حل مسأله^۵، ادراک^۶ و زبان^۷

به طور کلی تخمین خسارات ناشی از عملیات رزمی را می‌توان به دو دسته تقسیم نمود: خسارت‌های مستقیم شامل آن دسته از امکانات و تجهیزات و ظرفیت‌هایی است که بر اثر ضربه‌های مستقیم دشمن به نابودی کشیده شده است و شامل خسارت‌های وارده به ساختمان‌ها و تأسیسات، ماشین‌آلات، تجهیزات، مواد و کالا و غیره می‌باشد، و خسارت‌های غیرمستقیم که شامل امکانات و تجهیزات و ظرفیت‌هایی است که در صورت عدم بروز جنگ می‌توانست مورد بهره‌برداری قرار گیرد [۲]. در این مقاله هدف از ارزیابی و تخمین خسارت، همان خسارت‌های مستقیم می‌باشد که البته با محدودیت‌هایی مانند نبود داده‌های مناسب و کافی (تصاویر دریافتی قبل و بعد از عملیات رزمی) روبه‌رو می‌باشد.

دیان و همکاران (۲۰۱۵) [۸] الگوریتمی برای ارزیابی خسارات نبرد^{۱۰} با استفاده از تشخیص تغییرات در تصاویر رادار روزنه مصنوعی^{۱۱} (SAR) ارائه کردند. با توجه به دشواری تشخیص دقیق مناطق مورد حمله و غیر حمله در یک صحنه بزرگ SAR، تصاویر SAR به قسمت‌های کوچکی تقسیم می‌شوند و ویژگی‌های SIFT برای هر یک از آن‌ها استخراج می‌شود. با توجه به ویژگی‌های معنایی دو تصویر SAR در قبل و بعد از حمله محدوده‌های تغییر یافته، تشخیص و بررسی می‌شوند.

هان و جیان (۲۰۱۴) [۹] در مقاله خود به کاربرد شبکه بیزی در ارزیابی خسارت پرداخته‌اند. در این مقاله، تحلیل‌هایی بر روی شاخص‌های خسارت و متغیرهای حملات هوا به زمین، صورت می‌گیرد. با توجه به احتمال شرطی و محاسبات مبهم و تجربه کارشناسان، مدل شبکه بیزی ایجاد می‌شود. با پیچیدگی بیشتر محیط و افزایش پارامترهای نبرد، منابع اطلاعات، دارای عدم قطعیت بیشتری می‌شوند و بنابراین ارزیابی خسارت پیچیده‌تر خواهد بود و متغیرها؛ غیر خطی و غیر قطعی خواهند بود.

در روش پاول (۱۹۹۷) [۱۰] هدف، معرفی و توسعه مدل‌های تحلیلی احتمالاتی برای ارزیابی خسارت نبرد می‌باشد. به‌رغم دقت جنگ‌افزارهای جدید و دستگاه‌های ارتباطی و حسگرها، شلیک انجام شده به سمت هدف، گاهی اوقات از دست می‌رود یا آسیب جزئی وارد می‌آورد. در نتیجه توالی بیش از یک شلیک ممکن است مستقیماً به سمت هدف خاصی هدایت شود و بنابراین احتمال کشتار را افزایش می‌دهد. نقش ارزیابی خسارت در اینجا، بررسی این که آیا شلیک‌های بیشتر ضروری هستند یا نه، می‌باشد.

مهر (۱۹۹۹) [۱۱] یک فرآیند خودکار ارزیابی خسارت برای ارزیابی میزان آسیب نبرد، براساس تصاویر دریافتی از رادار لیزر در قبل و بعد از عملیات، معرفی نمود. این فرآیند شامل سه بخش است، بخش ثبت تصویر، بخش جداسازی خسارت، و بخش ارزیابی خسارت. در بخش ثبت تصویر، تصاویر دریافتی قبل و بعد از حمله به یکدیگر منطبق می‌شوند. در بخش جداسازی خسارت، تغییرات بین دو مجموعه تصاویر با موقعیت‌های جدا شده واقعی خسارت دیده مقایسه می‌شود. محل تقریبی آسیب، چنان چه وجود داشته باشد به بخش ارزیابی خسارت جهت تعیین وسعت خسارت با استفاده از روش رشد منطقه ارسال می‌شود.

۳- تشخیص اشیاء توسط یادگیری عمیق

منظور از اشیاء در تصاویر صحنه نبرد در واقع همان اهداف مورد نظر اپراتور سامانه ارزیابی خسارت می‌باشد. در این مقاله از روش شبکه‌های عصبی کانولوشن [۱۲] در یادگیری عمیق جهت تشخیص اشیاء مورد نظر (اهداف نظامی) استفاده می‌شود. یادگیری عمیق

متمرکز شده است [۱]. شبکه‌های عصبی از جمله الگوریتم‌های یادگیری ماشین در هوش مصنوعی می‌باشد که امروزه در اکثر پروژه‌های هوش مصنوعی و یادگیری ماشین از آن استفاده می‌شود. یادگیری عمیق^۴ در واقع شیوه مدرن یادگیری ماشین است که از شبکه‌های عصبی عمیق استفاده می‌کند. در این تحقیق نیز از روش یادگیری عمیق برای انجام عمل تشخیص اشیاء در تصاویر صحنه نبرد استفاده خواهیم کرد.

در روش پیشنهادی ما نیز تخمین خسارت نیاز به تشخیص و شناسایی اشیاء مورد نظر در تصاویر صحنه نبرد است. به عبارت دیگر با استفاده از تصاویری که از صحنه نبرد در قبل و بعد از عملیات رزمی دریافت می‌شود می‌توانیم عملیات ارزیابی خسارت را انجام دهیم.

۲- پیشینه تحقیق

پاجورجیو و ارن (۱۹۹۸) [۳] از یک چارچوب عمومی برای تشخیص اشیاء در تصاویر ثابت استفاده کردند. در این مقاله تشخیص اشیاء بر اساس نمایش طول موج و معیارهای آماری تحلیل شده است. در این روش، یادگیری یک کلاس شی از نظر یک زیرمجموعه‌ی کامل توابع مبتنی بر موجک، صورت می‌پذیرد و این کلاس شی به‌عنوان یک ورودی به طبقه‌بندی کننده‌ی ماشین بردار پشتیبان [۴] ارسال می‌شود.

در روش تورالبا و همکاران (۲۰۱۰) [۵] سیستم تشخیص اشیاء تنها می‌تواند تعداد محدودی از اشیاء را تشخیص دهد. تصویر ورودی در یک نمایش مناسب، به تصاویر مجموعه آموزشی که همگی برچسب خورده‌اند منطبق می‌شود. در اینجا یک مدل احتمالاتی برای انتقال برچسب‌ها از مجموعه بازبازی شده به تصویر ورودی ساخته می‌شود. رن و همکاران (۲۰۱۷) [۶] یک شبکه پیشنهاد منطقه‌ای (RPN)^۹ که ویژگی‌های کانولوشن کامل تصویر را با شبکه‌ی تشخیص به اشتراک می‌گذارد، ارائه نمودند و در نتیجه امکان پیشنهاد تقریباً بدون هزینه را فراهم کرده‌اند. به‌طور خلاصه RPN یک شبکه کاملاً کانولوشن است که به‌صورت هم‌زمان مرزهای شی را پیش‌بینی می‌کند. روش شبکه پیشنهاد منطقه‌ای به‌طور معمول بر ویژگی‌های ارزان و طرح‌های استنتاج اقتصادی تأکید دارد.

لیو و همکاران (۲۰۱۶) [۷] روشی برای تشخیص شی با استفاده از یک شبکه عصبی عمیق ارائه نموده‌اند. این روش فضای بیرونی جعبه‌های محاطی را به مجموعه‌ای از جعبه‌های پیش‌فرض با نسبت ابعاد و مقیاس در هر موقعیت نقشه‌ی ویژگی، گسسته‌سازی می‌کند. در زمان پیش‌بینی؛ شبکه، امتیازات را برای حضور هر دسته شی در هر جعبه‌ی پیش‌فرض و ایجاد تطبیق به هر جعبه برای انطباق بهتر شکل شی، تولید می‌کند.

به این API ها می‌توانید به سایت <https://github.com/tensorflow/models> مراجعه نمایید. اولین مرحله در فرایند تشخیص اشیاء، تهیه مجموعه داده‌ها یا تصاویر می‌باشد. تصاویر فراهم شده باید دارای کیفیت مناسب از نظر تباین، روشنایی و غیره باشند و همچنین باید شامل اشیاء مورد نظر ما برای تشخیص باشند. اشیاء مورد نظر برای تشخیص شامل ۲ کلاس تانک‌های جنگی و چادرهای نظامی هستند. در واقع شبکه عصبی بر روی این کلاس‌ها آموزش خواهد دید.

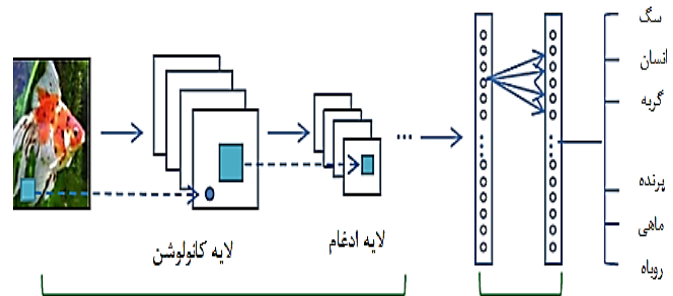
از آنجا که مجموعه داده‌ای که شامل این تصاویر باشد در بستر اینترنت وجود ندارد ما مجبور به جمع‌آوری داده‌های مورد نیاز توسط تصاویر موجود در سایت گوگل شده‌ایم.^{۱۷} مرحله بعدی برچسب زدن این داده‌ها توسط ابزارهای موجود است. این عمل به‌نوعی یک تعریف برای تصویر به حساب می‌آید، یعنی توسط این عمل مشخصات شی درون تصویر و همچنین مشخصات تصویر اصلی استخراج می‌گردد و در قالب فایل CSV^{۱۸} ذخیره می‌گردد. سپس فایل به دست آمده را در قالبی به نام TFRecords^{۱۹} به کتابخانه تنسورفلو می‌دهیم. از این قالب در تنسورفلو برای تغذیه داده‌ها به شبکه استفاده می‌شود.

بعد از انجام تمامی مراحل ذکر شده اکنون باید مدل خود را آموزش دهیم و برای این عمل می‌توان از مدل‌های از پیش آموزش داده شده برای تشخیص شی استفاده کرد. پیش آموزش به معنای آماده‌سازی و مقدار اولیه دادن به شبکه با پارامترهای از قبل آموزش داده شده به جای پارامترهای تصادفی است [۱۳]. با استفاده از مدل‌های از پیش آموزش داده شده، فرآیند آموزش شبکه می‌تواند بسیار سریع‌تر انجام شود و همچنین در مواقعی که اطلاعات و داده‌های مورد نیاز مانند تصاویر کم باشند از این روش استفاده می‌شود. برای این منظور مدل‌های متعددی از قبل تولید شده‌اند مانند، AlexNet [۱۴]، GoogleNet [۱۵] و MobileNet [۱۶]. هرکدام از این مدل‌ها خصوصیات و ویژگی‌های خاص خود را دارا هستند. به علت سادگی و همچنین سرعت بالاتر، در این مقاله از مدل MobileNet استفاده می‌کنیم. این مدل را می‌توانید از وبسایت زیر دریافت کنید:

http://download.tensorflow.org/models/object_detection/ssd_MobileNet_v1_coco_2018_01_28.tar.gz

شبکه MobileNet از لایه‌های متنوعی تشکیل شده است و عمده‌ترین لایه‌ها در این شبکه لایه‌های کانولوشن هستند که وظیفه‌ی استخراج ویژگی‌ها از تصاویر را بر عهده دارند. در جدول ۱ می‌توانید ساختار این شبکه را مشاهده نمایید. در پیاده‌سازی این مقاله نسخه‌ای از شبکه MobileNet استفاده می‌شود که قبلاً بر

زیرمجموعه‌ای از یادگیری ماشین است به‌طوری که از شبکه‌های عصبی عمیق استفاده می‌کند. شبکه‌های عصبی کانولوشن برای پردازش اطلاعات تصویر ابداع شده‌اند، از این رو باید ویژگی‌های تصاویر و پارامترها توسط این شبکه‌ها استخراج شده و در امر یادگیری مورد استفاده قرار گیرند. شکل ۱ نمایی از ساختار شبکه‌های عصبی کانولوشن را نمایش می‌دهد.



شکل ۱: نمایی از ساختار شبکه عصبی کانولوشن

به‌طور خلاصه، شبکه‌های کانولوشنی از اتصالاتی به نام شبکه‌ی استخراج ویژگی و شبکه‌ی طبقه‌بند یا دسته‌بند، تشکیل شده‌اند و شبکه استخراج ویژگی از لایه‌های کانولوشنی و ادغام^{۱۲} تشکیل شده است. عملیات کانولوشن توسط فیلترهایی برای استخراج ویژگی‌ها انجام می‌شود. فیلترها به‌صورت مجموعه‌ای از اعداد مانند ماتریس‌ها در کنار یکدیگر قرار گرفته‌اند. عملیات کانولوشن با حرکت این فیلترها بر روی تصویر و ضرب داخلی فیلتر و بخشی از تصویر که فیلتر بر روی آن قرار دارد انجام می‌شود. وظیفه لایه ادغام تغییر اندازه تصویر می‌باشد و در واقع اندازه آن را کاهش می‌دهد. بخش دیگری نیز ویژگی‌های به دست آمده را طبقه‌بندی می‌کند که به آن لایه تماماً متصل^{۱۳} می‌گویند. لایه تماماً متصل درنهایت ویژگی‌های دوبعدی را به یک بردار یک بعدی جهت نمایش نام دسته تبدیل می‌کند [۶].

۳-۱-۳ مراحل پیاده‌سازی تشخیص اشیاء

در این مقاله از کتابخانه متن‌باز تنسورفلو، که توسط گروه تحقیقاتی گوگل^{۱۴} ابداع شده است، برای پیاده‌سازی سامانه تشخیص اشیاء در تصاویر صفحه نبرد استفاده می‌کنیم. تنسورفلو یک کتابخانه نرم‌افزاری متن‌باز برای محاسبات عددی با کارایی بالا است.^{۱۵} استفاده از امکانات یک کتابخانه یا بسته‌ی نرم‌افزاری، مستلزم استفاده از رابط‌های برنامه‌نویسی کاربردی (API)^{۱۶} آن است. تنسورفلو، API های متفاوتی را فراهم ساخته است یکی از API های تنسورفلو، عملیات تشخیص اشیاء را انجام می‌دهد. برای دسترسی

میان تصاویر، در جهت انطباق پیکسل‌های مشابه در دو تصویر قبل و بعد از نبرد استفاده می‌شود.

جدول ۱: معماری شبکه MobileNet

نوع لایه	اندازه فیلتر	اندازه ورودی
Conv / s2	$3 \times 3 \times 3 \times 32$	$224 \times 224 \times 3$
Conv dw / s1	$3 \times 3 \times 32 \text{ dw}$	$112 \times 112 \times 32$
Conv / s1	$1 \times 1 \times 32 \times 64$	$112 \times 112 \times 32$
Conv dw / s2	$3 \times 3 \times 64 \text{ dw}$	$112 \times 112 \times 64$
Conv / s1	$1 \times 1 \times 64 \times 128$	$56 \times 56 \times 64$
Conv dw / s1	$3 \times 3 \times 128 \text{ dw}$	$56 \times 56 \times 128$
Conv / s1	$1 \times 1 \times 128 \times 128$	$56 \times 56 \times 128$
Conv dw / s2	$3 \times 3 \times 128 \text{ dw}$	$56 \times 56 \times 128$
Conv / s1	$1 \times 1 \times 128 \times 256$	$28 \times 28 \times 128$
Conv dw / s1	$3 \times 3 \times 256 \text{ dw}$	$28 \times 28 \times 256$
Conv / s1	$1 \times 1 \times 256 \times 256$	$28 \times 28 \times 256$
Conv dw / s2	$3 \times 3 \times 256 \text{ dw}$	$28 \times 28 \times 256$
Conv / s1	$1 \times 1 \times 256 \times 512$	$14 \times 14 \times 256$
$5 \times$ Conv dw / s1	$3 \times 3 \times 512 \text{ dw}$	$14 \times 14 \times 512$
Conv / s1	$1 \times 1 \times 512 \times 512$	$14 \times 14 \times 512$
Conv dw / s2	$3 \times 3 \times 512 \text{ dw}$	$14 \times 14 \times 512$
Conv / s1	$1 \times 1 \times 512 \times 1024$	$7 \times 7 \times 512$
Conv dw / s2	$3 \times 3 \times 1024 \text{ dw}$	$7 \times 7 \times 1024$
Conv / s1	$1 \times 1 \times 1024 \times 1024$	$7 \times 7 \times 1024$
Avg Pool / s1	Pool 7×7	$7 \times 7 \times 1024$
FC / s1	1024×1000	$1 \times 1 \times 1024$
Softmax / s1	Classifier	$1 \times 1 \times 1000$

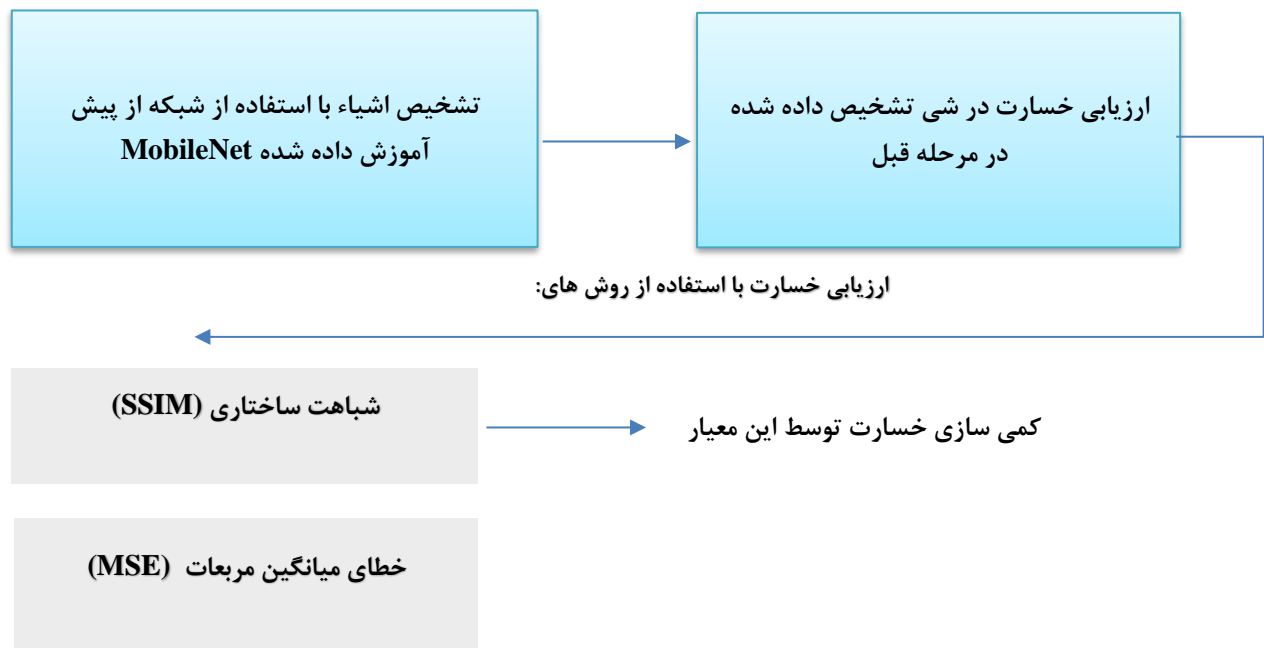
روی مجموعه داده COCO آموزش داده شده است [۱۷]. علت استفاده از این نوع مجموعه داده، تنوع و غنی بودن تصاویر داخل این مجموعه است. در ادامه لایه انتهایی این شبکه به نام Softmax layer [۱۸] وظیفه انتساب احتمال به هر یک از کلاس‌های اشیاء تشخیص داده شده توسط شبکه را بر عهده دارد. در این لایه بیشترین احتمال متعلق به کلاسی است که به درستی تشخیص داده شده است.

پس از اینکه اشیاء مورد نظر توسط شبکه تشخیص داده شدند نوبت به عملیات ارزیابی خسارت می‌رسد که در بخش بعدی به تفصیل مورد بررسی قرار خواهد گرفت. روش پیشنهادی ما در شکل ۲ نمایش داده شده است.

۴- ارزیابی خسارت

در این مقاله، پیاده‌سازی عملیات ارزیابی خسارت، توسط انطباق و ثبت تصاویر قبل و بعد از نبرد صورت می‌گیرد. هدف از ثبت تصویر، تطبیق دادن دو یا چند تصویر که با زوایای دید مختلف یا در فواصل مختلف و یا در زمان‌های مختلف از یک صحنه تهیه شده‌اند، می‌باشد.

در این مقاله، جهت انجام فرآیند ارزیابی خسارت از معیارهای خطای میانگین مربعات (MSE)^۱، شباهت ساختاری (SSIM) [۱۹] و همچنین آستانه گذاری^۲، جهت سنجش میزان شباهت و تغییرات



شکل ۲: مراحل روش پیشنهادی

در نهایت، هر سه قسمت در یک فرمول نهایی خلاصه می‌شوند:

$$S(x, y) = f(l(x, y), c(x, y), s(x, y)) \quad (4)$$

به طوری که برای قسمت روشنایی تصویر (پارامتر اول) داریم:

$$l(x, y) = \frac{2\mu_x \mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \quad (5)$$

ثابت C_1 برای جلوگیری از ناپایداری محاسباتی هنگامی که $\mu_x^2 + \mu_y^2$ به مقدار صفر بسیار نزدیک باشد استفاده می‌شود و به صورت زیر تعیین می‌شود.

$$C_1 = (K_1 L)^2 \quad (6)$$

که در آن L محدوده پویای مقادیر پیکسل (۲۵۵) برای تصاویر خاکستری ۸ بیت، و $K_1 \leq 1$ مقدار ثابت کوچکی می‌باشد. برای قسمت تباین تصویر داریم:

$$c(x, y) = \frac{2\sigma_x \sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \quad (7)$$

مانند رابطه (۶)، مقدار ثابت C_2 برابر با $(K_2 L)^2$ است. و $K_2 \leq 1$ است. برای قسمت مقایسه ساختاری تصویر داریم:

$$s(X, Y) = \frac{\sigma_{xy} + C_3}{\sigma_x \sigma_y + C_3} \quad (8)$$

ثابت C_3 مانند روابط قبلی محاسبه می‌شود و در شکل گسسته، σ_{xy} می‌تواند به صورت رابطه زیر تخمین زده شود:

$$\sigma_{xy} = \frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)(y_i - \mu_y) \quad (9)$$

فرمول نهایی برای روش شباهت ساختاری به صورت زیر بیان می‌شود:

$$SSIM(x, y) = \frac{(2\mu_x \mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (10)$$

پارامترهای موجود در این فرمول شامل (x, y) مکان پنجره $N * N$ در هر تصویر می‌باشد. μ_x, μ_y میانگین شدت‌های روشنایی پیکسل‌ها در جهت x و y ، و σ_x^2, σ_y^2 واریانس شدت‌های روشنایی پیکسل‌ها در جهت x و y همراه با کوواریانس (σ_{xy}) آن‌ها است. برای انجام عمل ارزیابی خسارت توسط پردازش تصاویر قبل و بعد از نبرد، ابتدا اشیاء مورد نظر ما در تصویر تشخیص داده می‌شوند، این تشخیص می‌تواند در کنترل تجهیزات نظامی و همچنین تصمیم‌گیری و ارائه راهکار و ارزیابی و برآورد پارامترهای مهم در نبرد، مؤثر باشد. برای انجام عمل ارزیابی خسارت، ترجیح

روش شباهت ساختاری، تغییرات را در اطلاعات ساختاری می‌سنجد. اما روش خطای میانگین مربعات، خطاهای دریافتی را تخمین می‌زند. روش شباهت ساختاری، دو تصویر را با استفاده از پنجره‌هایی مقایسه می‌کند، در واقع نمونه‌های کوچک‌تری از تصاویر را توسط پنجره‌هایی با یکدیگر مقایسه می‌کند. در روش خطای میانگین مربعات، کل تصاویر برای مقایسه با یکدیگر دریافت می‌شوند. مشکل عمده در روش خطای میانگین مربعات، این است که فاصله‌های زیاد بین شدت‌های پیکسل، لزوماً به این معنی نیست که محتویات تصاویر تفاوت‌های زیادی دارند. در این مقاله برای کنترل بیشتر و تصمیم‌گیری بهتر فرمانده، از هر دو روش استفاده نموده‌ایم.

اساس کلی روش خطای میانگین مربعات مطابق فرمول (۱) می‌باشد:

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - K(i, j)]^2 \quad (1)$$

که در آن m و n ابعاد تصاویر هستند و $I(i, j)$ و $K(i, k)$ به ترتیب، تصاویر اول و دوم با درایه‌های i و j می‌باشند. در روش شباهت ساختاری، اندازه‌گیری شباهت به سه قسمت تقسیم می‌شود: روشنایی^{۲۲}، تباین^{۲۴} و ساختار^{۲۵}.

در ابتدا روشنایی هر تصویر ارزیابی می‌شود. از رابطه (۲) برای تخمین میانگین شدت روشنایی استفاده می‌شود:

$$\mu_x = \frac{1}{N} \sum_{i=1}^N x_i \quad (2)$$

که در آن x_i مقادیر پیکسل‌های تصویر و N تعداد پیکسل‌های تصویر x هستند. مطابق رابطه (۲) میانگین شدت روشنایی برای تصویر y نیز محاسبه می‌شود (μ_y) . تابع مقایسه روشنایی $l(x, y)$ تابعی از μ_x و μ_y است. در مرحله بعد، میانگین شدت روشنایی را از تصویر حذف می‌کنیم $(x - \mu_x)$.

از انحراف معیار^{۲۶} (ریشه مربع واریانس)، به عنوان تخمینی از تباین تصویر، استفاده می‌کنیم. مطابق فرمول:

$$\sigma_x = \left(\frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)^2 \right)^{\frac{1}{2}} \quad (3)$$

تابع تباین به صورت عبارت $c(x, y)$ بیان می‌شود و σ_x و σ_y را مقایسه می‌کند. سپس تصویر توسط انحراف معیار خودش نرمال‌سازی می‌شود.

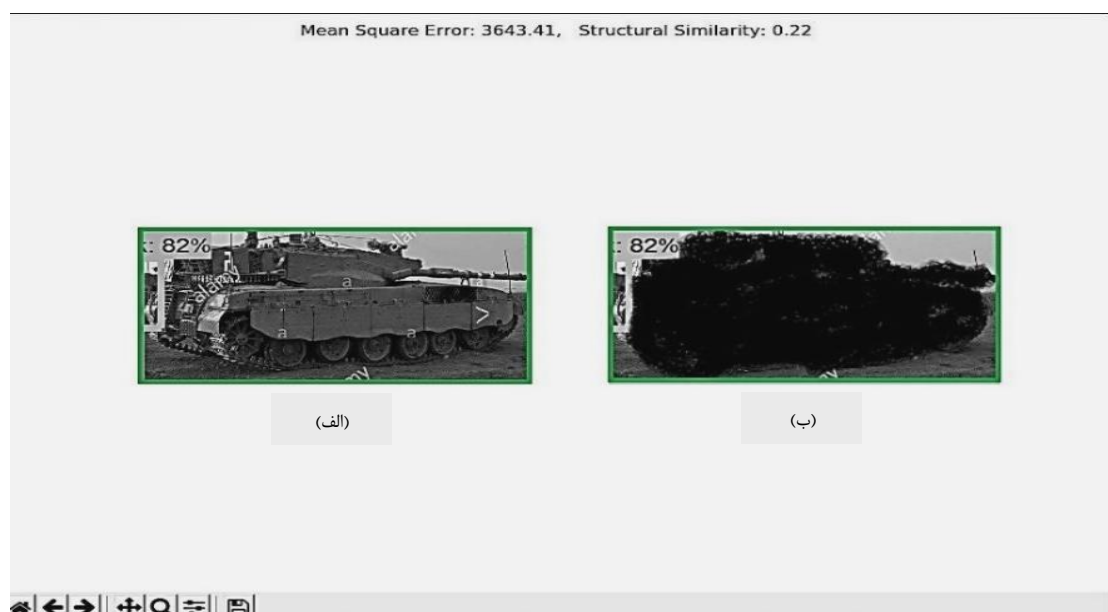
مقایسه ساختاری $s(x, y)$ ، بر روی تصاویر نرمال‌سازی شده $(x - \mu_x)$ و $(y - \mu_y)$ انجام می‌شود.

بر این است که ابتدا اشیاء مهم (از نظر اپراتور) در تصویر تشخیص داده شوند و سپس خسارت وارده به همان اشیاء تخمین زده شود. در یک تصویر ممکن است اشیایی در تصویر موجود باشد که به دست آوردن خسارت وارده به آن‌ها برای ما اهمیتی نداشته باشد و به همین دلیل کار تشخیص شیء به ما کمک می‌کند تا قسمت‌های مهم تصویر را استخراج کرده و ارزیابی خسارت را فقط بر روی همان قسمت تشخیص داده شده انجام دهیم. در مقابل، مواردی نیز پیش می‌آید که اپراتور نیاز دارد تمام تصویر را از لحاظ خسارت ارزیابی کند. در این مورد نیازی به تشخیص اشیاء نبوده و می‌توان تمام تصویر را از نظر خسارت وارده تخمین و ارزیابی نمود. به عنوان مثال در شکل ۳ ارزیابی خسارت فقط درون کادر شیء تشخیص داده شده توسط الگوریتم انجام شده است. همان‌طور که در تصویر نمایان است مقدار شباهت ساختاری (SSIM) برابر ۰.۲۲ و مقدار خطای میانگین مربعات (MSE) برابر با ۳۶۴۳.۴۱ است. طبق ارزیابی الگوریتم، مقدار شباهت ساختاری هرچه به عدد ۱ نزدیک‌تر باشد میزان تغییرات کمتر و خسارت کمتری خواهیم داشت و هر چه این مقدار به ۰ نزدیک‌تر شود میزان شباهت کمتر و در مقابل خسارت بیشتری به تجهیزات وارد گشته است. در توضیح معیار خطای میانگین مربعات، هرچه مقدار آن به صفر نزدیک‌تر باشد به اصطلاح خطای کمتر و شباهت بیشتری بین تصاویر وجود دارد و در نهایت خسارت کمتری خواهیم داشت و هرچه این عدد بیشتر شود شباهت کمتری وجود دارد و تغییرات بیشتری مشاهده می‌شود، پس در مقابل، خسارت بیشتری نیز به تجهیزات وارد گشته است. در شکل ۳ ابتدا تانک جنگی توسط شبکه عصبی

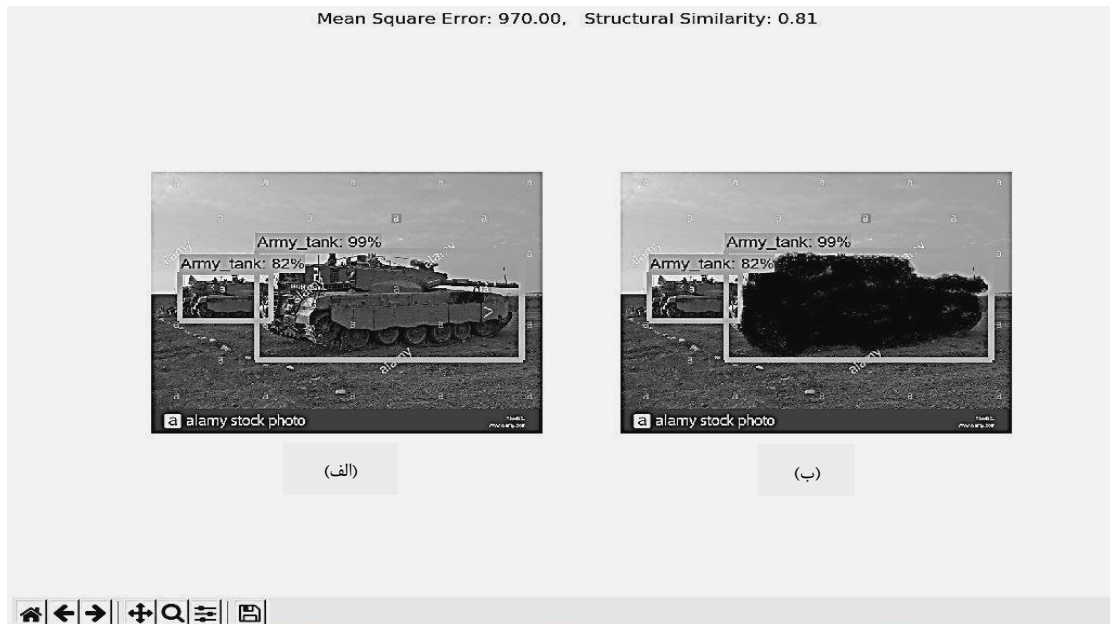
MobileNet تشخیص داده می‌شود و کادری سبز رنگ به دور آن ترسیم می‌کند و در ادامه برای سنجش میزان خسارت در همان شیء تشخیص داده شده توسط شبکه، باید شیء مربوطه را به صورت دستی یا به صورت خودکار توسط الگوریتم ببریم که ما در اینجا به صورت دستی این عمل را انجام داده‌ایم. طبق معیار شباهت ساختاری نمایش داده شده (در بالای شکل ۳)، به این نتیجه می‌رسیم که تانک مورد نظر به‌طور تقریباً کامل از بین رفته است و در واقع حدود ۸۰ درصد این تانک تخریب شده است شکل ۳ (ب). همچنین معیار خطای میانگین مربعات (در بالای شکل ۳) نشان می‌دهد که تغییرات بسیاری صورت گرفته و تانک مورد نظر به‌طور تقریباً کامل تخریب شده است. در پیاده‌سازی این مقاله با توجه به کمبود تصاویر مورد نیاز، تمامی تغییرات آسیب‌های وارده به تانک‌ها و دیگر تجهیزات نظامی در تصویر؛ در زمان بعد از نبرد، شبیه‌سازی شده‌اند. در ادامه از روش SSIM جهت کمی‌سازی میزان خسارت از نظر وسعت، استفاده خواهیم نمود (جدول ۲).

جدول ۲: رده بندی میزان خسارت

میزان خسارت	شاخص خسارت	سطح خسارت
کم	$0.75 < SSIM \leq 0.99$	۱
متوسط	$0.50 < SSIM \leq 0.75$	۲
شدید	$0.25 < SSIM \leq 0.50$	۳
کامل	$0 < SSIM \leq 0.25$	۴



شکل ۳: تشخیص تانک و ارزیابی خسارت در داخل کادر سبز رنگ. (الف) تصویر تانک سالم قبل از نبرد. (ب) تصویر تانک تخریبی بعد از نبرد.



شکل ۴: ارزیابی خسارت در تمامی تصویر. (الف) تصویر تانک سالم قبل از نبرد. (ب) تصویر تانک تخریب شده بعد از نبرد

به وجود آمده در مرحله بعد از نبرد را توسط معیار حدآستانه نمایان نموده و کادری قرمز رنگ به دور تمامی تغییرات ترسیم می‌نماید و همچنین تغییرات را به صورت تصاویر سیاه و سفید نمایش می‌دهد تا به خوبی بتوان موقعیت و میزان آسیب‌دیدگی را مشاهده نمود. به عنوان مثال در شکل ۵ تصویری از چادرهای نظامی که به صورت شبیه سازی شده، بخشی از آن‌ها بر اثر سوختگی از بین رفته است نمایش داده می‌شود.

۵- آزمایش‌ها

همان‌طور که قبلاً بیان شد برای تشخیص اشیاء در تصاویر از روش شبکه‌های عصبی کانولوشن و کتابخانه متن‌باز تنسورفلو استفاده نموده‌ایم. در این بین آزمایش‌هایی مانند تشخیص اشیاء، ارزیابی خسارت، سنجش میزان خطای شبکه، مدت زمان موردنیاز برای تکمیل فرآیند آموزش شبکه انجام شده است که در ادامه به تحلیل هر یک خواهیم پرداخت. در پیاده‌سازی و آزمایش روش پیشنهادی از سیستم کامپیوتری با پردازنده اینتل Corei7-3630QM و ۶ گیگابایت حافظه اصلی استفاده شده است. در شکل ۶ عملیات تشخیص چادر نظامی نمایش داده شده است.

همان‌طور که در شکل ۳ مشاهده می‌شود مقدار SSIM برابر با ۰.۲۲ است که این میزان در سطح ۴ جدول رده‌بندی خسارت قرار دارد و نشان دهنده تخریب کامل است.

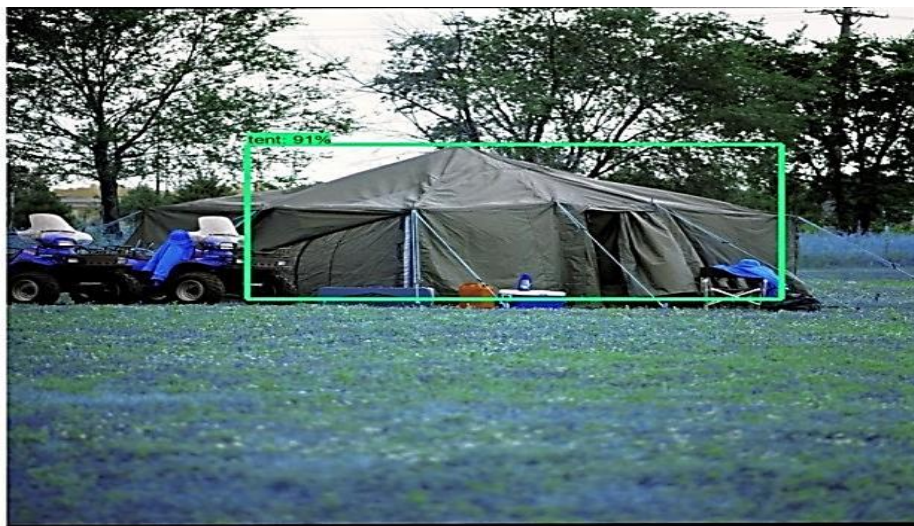
در شکل ۴ ارزیابی خسارت در تمام تصویر انجام شده و تنها داخل کادر ارزیابی نمی‌شود. همان‌طور که مشاهده می‌شود مقدار شباهت ساختاری برابر ۰.۸۱ و خطای میانگین مربعات برابر ۹۷۰.۰۰ است و این در حالی است که الگوریتم، تانک آسیب‌دیده را نسبت به تمامی تصویر می‌سنجد، در حالی که به تجهیزات در مکان‌های دیگر تصویر خسارتی وارد نشده و تنها به یکی از تانک‌های درون کادر آسیب رسیده است. پس در این مورد بهتر است که ارزیابی خسارت تنها در داخل کادر صورت پذیرد تا نتایج حاصل دقیق‌تر باشند

بعد از به دست آوردن میزان شباهت و تغییرات تصاویر و همچنین موقعیت تانک‌های جنگی در تصاویر صحنه نبرد، جهت ارزیابی خسارت از لحاظ مالی به کارشناس خبره نظامی نیاز است که بتواند با استفاده از اطلاعاتی که برنامه در اختیار او قرار می‌دهد، ارزیابی قیمت تجهیزات را انجام دهد و در نهایت خسارت دقیق آن را برحسب واحد معینی برآورد کند.

برای کمک به اپراتور در جهت تصمیم‌گیری بهتر و همچنین کنترل بیشتر بر روی عملیات ارزیابی خسارت، ویژگی دیگری نیز به سامانه ارزیابی خسارت افزوده شده است. این ویژگی تغییرات



شکل ۵: نمایش تغییرات و نقاط آسیب دیده بعد از نبرد توسط خطوط قرمز رنگ (راست تصویر) و تصاویر سیاه و سفید (چپ تصویر)



شکل ۶: چادر نظامی تشخیص داده شده توسط شبکه عصبی کانولوشن (MobileNet)

می‌یابد. حتی با وجود سیستم‌های قدرتمند باز هم زمان مورد نیاز برای آموزش شبکه از بسیاری الگوریتم‌های متداول دیگر که از یادگیری عمیق استفاده نمی‌کنند، بیشتر است. با توجه به حجم داده‌های ورودی به شبکه و قدرت سیستم کامپیوتری، ممکن است آموزش شبکه حتی به چندین روز زمان نیاز داشته باشد. به‌عنوان مثال، در پیاده‌سازی روش پیشنهادی این زمان تقریباً ۳ روز به طول انجامید. در نمودار ۱ تغییرات میزان خطای شبکه نشان داده شده است. همان‌طور که در نمودار مشاهده می‌شود میزان خطای شبکه به‌مرور در حال کاهش است.

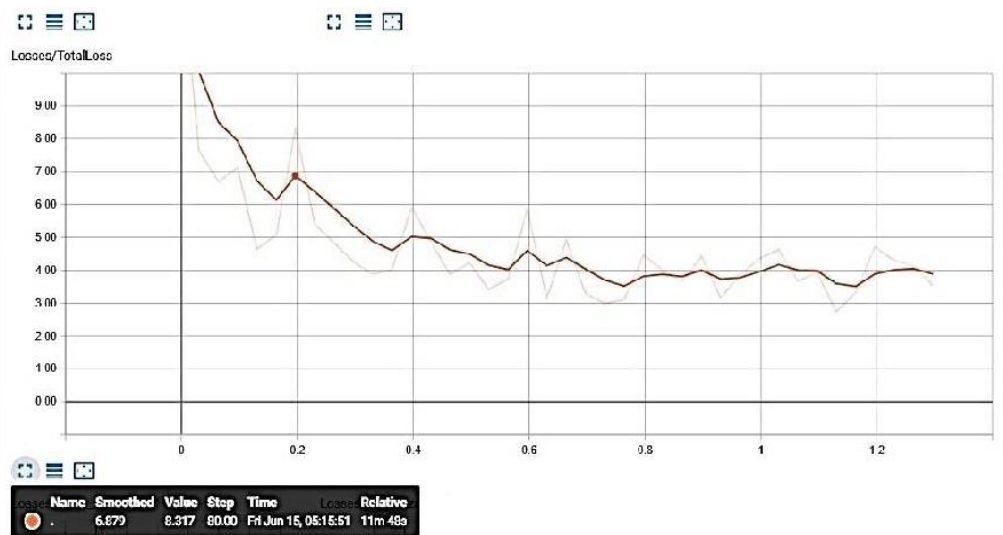
همان‌طور که در تصویر مشاهده می‌شود، چادر نظامی به احتمال ۹۱٪ تشخیص داده شده است و این نشان از دقت بالای تشخیص می‌باشد.

در شکل ۷ (الف) تصویر بازارچه را در قبل از وارد شدن خسارت نشان می‌دهد و در شکل ۷ (ب) محدوده بازارچه تقریباً کامل از بین رفته است و میزان شباهت ساختاری نشان می‌دهد که به حدود ۱/۴ بازارچه آسیبی نرسیده است که شامل ساختار اصلی بازارچه است. همچنین خطای میانگین مربعات عدد بالایی را نمایش می‌دهد که نشان از تغییرات زیاد در تصویر است.

همان‌طور که بیان شد، در یادگیری عمیق و شبکه‌های عصبی کانولوشن، با حجم بسیار زیادی از داده‌ها مواجه هستیم و به همین دلیل زمان مورد نیاز برای آموزش شبکه بسیار افزایش



شکل ۷: تصاویر دریافتی قبل و بعد از تخریب. (الف) قبل از تخریب (ب) بعد از تخریب



نمودار ۱: نمایش روند تغییرات خطای شبکه

۶- نتیجه گیری

در این مقاله ما از روش‌های هوش مصنوعی برای کمک به اپراتور یا فرمانده عملیات در جهت پیشبرد اهداف نظامی استفاده نموده‌ایم. به طوری که کار طاقت فرسای تعیین و مشاهده تغییرات به وجود آمده در صحنه نبرد در بعد از عملیات رزمی را به دوش سامانه ارزیابی خسارت نهاده و سعی کردیم تا حد ممکن از دخالت نیروی انسانی بکاهیم.

در سامانه ارزیابی خسارت از دو معیار شباهت ساختاری و خطای میانگین مربعات برای تعیین میزان شباهت تصاویر استفاده شده است.

این سامانه همچنین به اپراتور کمک می‌کند تا موقعیت اشیاء مورد نظر خود را در تصویر به دست آورده و تصمیمات مناسب را اتخاذ نماید. این سامانه برای عمل تشخیص اشیاء، محل آن‌ها را توسط ترسیم کادری به دور شی، در تصویر مشخص می‌کند. برای این کار با استفاده از مجموعه داده‌های آموزشی یک شبکه عصبی عمیق توسط مدل از پیش آموزش شده MobileNet آموزش داده شد که هر شی را به عنوان یک کلاس خاص تشخیص دهد. استفاده از این روش دقت تشخیص اشیاء را نسبت به روش‌هایی که نیاز است شبکه از ابتدا آموزش ببیند، افزایش می‌دهد. با توجه به حجم بالای ویژگی‌ها و پارامترهای استخراج شده از داده‌ها در

- [7] W. Liu et al., "SSD: Single shot multibox detector," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 9905 LNCS, pp. 21-37, 2016.
- [8] Z. De-yun, Z. Li-na, Z. Kai, and Z. Kun, "Battle damage assessment with change detection of SAR images," Control Conf. (CCC), 34th Chinese, pp. 193-197, 2015.
- [9] L. Chen-han and H. Jian, "The application of Bayesian network in battle damage assessment," Softw. Eng. Serv. Sci, pp. 529-532, 2014.
- [10] D. Paul, "Probability models for battle damage assessment (simple shoot-look-shoot and beyond)," Reports and Technical Reports, Monterey, California, pp. 1-26, 1997.
- [11] R. K. Mehraa, "Automatic Battle Damage Assessment based on Laser Radar Imagery," vol. 3707, no., pp. 210-221, 1999.
- [12] Y. Lecun, L. Bottou, Y. Bengio and P. Haffner, "Gradient-based learning applied to document recognition," in Proceedings of the IEEE, vol. 86, no. 11, pp. 2278-2324, Nov. 1998
- [13] Rawat, W. and Wang, Z., "Deep convolutional neural networks for image classification: A comprehensive review". Neural computation, 29(9), pp.2352-2449, 2017.
- [14] A. Krizhevsky, I. Sutskever, and H. Geoffrey E, "ImageNet Classification with Deep Convolutional Neural Networks," Adv. Neural Inf. Process, Syst. 25, pp. 1-9, 2012.
- [15] C. Szegedy, S. Reed, P. Sermanet, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1-12, 2014.
- [16] A. G. Howard et al., "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," Computer Vision and Pattern Recognition, pp. 1-9, 2017.
- [17] Lin TY. et al. "Microsoft COCO: Common Objects in Context. "In: Fleet D., Pajdla T., Schiele B., Tuytelaars T. (eds) Computer Vision – ECCV 2014. ECCV 2014. Lecture Notes in Computer Science, vol 8693, 2014.
- [18] Goodfellow, Ian; Bengio, Yoshua; Courville, Aaron, "6.2.2.3 Softmax Units for Multinoulli Output Distributions". Deep Learning. MIT Press. pp. 180-184. ISBN 978-0-26203561-3. 2016.
- [19] 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, 2004.

شبکه عصبی عمیق، سرعت تجزیه و تحلیل داده‌ها تا حدی زیادی کاهش یافت.

پایه‌سازی این پروژه با توجه به محدودیت‌هایی مانند کمبود داده‌های مورد نیاز (تصاویر دریافتی از منطقه نبرد، قبل و بعد از رخداد خسارت) و همچنین نبود سامانه‌های کامپیوتری با قدرت پردازش بالا انجام شده است. با وجود تمامی این مشکلات سامانه‌ای طراحی گردیده است که می‌تواند برخی اشیاء موجود در تصاویر نبرد را تشخیص دهد و با توجه به اشیاء موجود در تصویر و حتی در بعضی مواقع بدون در نظر گرفتن عملیات تشخیص شی، عمل ارزیابی خسارت را انجام دهد.

مراجع

- [1] D. L. Poole and A. K. Mackworth, "Artificial intelligence: Foundations of computational agents," Cambridge University Press, vol. 9780521519, 2010.
- [۲] دژ پسند، فرهاد، و حمیدرضا رئوفی، "اقتصاد ایران در دوران جنگ تحمیلی." فصلنامه تخصصی مطالعات دفاع مقدس، صفحه ۱۵۰-۱۴۷، ۱۳۸۷.
- [3] C. P. Papageorgiou and M. Oren, "A general framework for object detection," Comput Vision, IEEE Int. Conf, vol. 0, no. pp. 555-562, 1998.
- [4] C. Cortes, V. Vapnik, "Support Vector Networks," Mach. Learn., vol. 20, no. 3, pp. 273-297, 1995.
- [5] a. Torralba, A. Oliva, and W. T. Freeman, "Object recognition by scene alignment," J. Vis, vol. 3, pp. 196-204, 2010.
- [6] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," IEEE Trans. Pattern Anal. Mach. Intell, vol. 39, no. 6, pp. 1440-1448, 2017.

زیر نویس‌ها:

¹⁵ <https://www.tensorflow.org>

¹⁶ Application Programming Interface

¹⁷ <https://www.pyimagesearch.com/2017/12/04/how-to-create-a-deep-learning-dataset-using-google-images/>

¹⁸ Comma Separated Values

¹⁹ Tensorflow Records

²⁰ Pre-Training

²¹ Mean Square Error

²² Thresholding

²³ Luminance

²⁴ Contrast

²⁵ Structure

²⁶ Standard Deviation

¹ Object Detection

² Artificial Intelligence

³ Learning

⁴ Reasoning

⁵ Problem Solving

⁶ Perception

⁷ Language

⁸ Deep Learning

⁹ Region Proposal Network

¹⁰ Battle Damage Assessment

¹¹ Synthetic Aperture Radar

¹² Pooling

¹³ Fully Connected Layer

¹⁴ Google