Maritime management science studies

Human Detection Using Thermal Cameras

Amirhosein zanganeh[1]

Abstract

purpose: Today, due to the existence of various threats to individuals and private and public places, all efforts are made to use the available technological advances to prevent accidents and protect people and their places. Today, camera-based surveillance systems are widely used to monitor and control the environment. Conventional cameras used in these systems in different weather conditions or during the night are challenging to identify and recognize objects and provide poor results. To increase security and surveillance in protected areas, it is very important to identify people and suspicious movements in different weather conditions 24 hours a day.

Methodology: This research is applicable in terms of purpose and the results of this study are presented using simulation - based methods.

Findings: With the advancement of technology and the widespread use of thermal images, we are witnessing a revolutionary process in the field of detecting and identifying objects. Thermal images allow us to obtain information about the amount of heat transfer of a surface and its shape. In this article, our goal is to extract and analyze useful information from thermal images to identify and diagnose humans using thermal cameras.

Conclusion: We have used the YOLO3 network to identify and recognize objects in thermal video. Based on the results, with mAP = 94.6, we have been able to identify and recognize humans in the output videos from the thermal camera.

Key words: Thermal image, human recognition, YOLO algorithm, deep neural network, deep learning.

1. Assistant Professor of Electrical Engineering, Shahid Sattari Aeronautical University, Tehran, Iran.
E-mail: **zangane.amirhosein@ssau.ac.ir**

Journal of maritime management sciences,2023,vol. 4,No 2,pp.99-110
Doi:https://doi.org/ 10.22034/MMR.2024.316658.1124
Article Type: Research-based Published by Faculty of Management and Marine Commissary
Received: May 5, 2023; Accepted: January 6,2023

فصلنامه مطالعات علوم مدیریت دریایی

تشخیص انسان با استفاده از دوربین‌های حرارتی

امیرحسین زنگنه[1*]

چکیده

هدف: امروزه با توجه به وجود تهدیدات متنوع برای برای افراد و اماکن خصوصی و دولتی تمام تلاش‌ها برای بهره‌برداری از پیشرفت‌های تکنولوژیک موجود برای جلوگیری از حوادث ناگوار و حفاظت از مردم و اماکن آنها انجام می‌شود. برای نظارت و کنترل محیط امروزه به طور گسترده از سیستمهای نظارتی مبتنی بر دوربین استفاده می‌شود. دوربین های معمولی مورد استفاده در این سیستمها در شرایط آب و هوایی مختلف و یا در طول شب برای شناسایی و تشخیص اشیاء با چالش روبرو هستند و نتایج ضعیفی ارائه می دهند. برای افزایش امنیت و نظارت در مناطق حفاظت شده، شناسایی افراد و تحرکات مشکوک در شرایط آب و هوایی مختلف و در طول ۲۴ ساعت شبانه روز بسیار مهم است.

روش شناسی: این تحقیق از نظر هدف کاربردی است و نتایج این پژوهش با استفاده از روش‌های مبتنی بر شبیه‌سازی ارائه شده است.

یافته‌ها: با پیشرفت فناوری و استفاده گسترده از تصاویر حرارتی، شاهد روندی تحولی در حوزه تشخیص و شناسایی اشیا هستیم. تصاویر حرارتی، امکان دست یابی به اطلاعاتی در مورد میزان انتقال حرارت یک سطح و شکل آن را برای ما فراهم می‌کنند. در این مقاله، هدف ما، استخراج و تحلیل اطلاعات مفید تصاویر حرارتی با هدف شناسایی و تشخیص انسان با استفاده از دوربینهای حرارتی است.

نتیجه‌گیری: ما برای شناسایی و تشخیص اشیاء در ویدیوی حرارتی از شبکه یولو۳ استفاده کرده‌ایم. براساس نتایج حاصل با ۹۴/۶ = mAP قادر به شناسایی و تشخیص انسان در ویدیوهای خروجی از دوربین حرارتی شده‌ایم.

واژه‌های کلیدی: تصویر حرارتی، تشخیص انسان، الگوریتم یولو، شبکه عصبی عمیق، یادگیری عمیق.

استناد: زنگنه, امیرحسین. (۱۴۰۲). تشخیص انسان با استفاده از دوربین‌های حرارتی .*مطالعات علوم مدیریت دریایی*,(۲)۴.

۱.استادیار، دانشکده مهندسی کامپیوتر دانشگاه هوایی شهید ستاری، تهران ، ایران **نویسندهٔ مسئول ایمیل:** **zangane.amirhosein@ssau.ac.ir**.

نوع مقاله: علمی و پژوهشی ناشر:دانشکده مدیریت و کمیسردریایی دانشگاه علوم دریایی امام خمینی(ره)

Doi:https://doi.org/ 10.22034/MMR.2024.316658.1124

تاریخ دریافت مقاله:۱۴۰۱/۱۱/۰۶؛ **تاریخ پذیرش مقاله:۱۴۰۲/۰۲/۱۵**

مقدمه

به دلیل تهدیدات تروریستی جهانی و مهاجرت غیرقانونی، نگرانی ها در مورد امنیت شهروندان تشدید شده است. برای جلوگیری از حوادث ناخواسته و محافظت از اماکن نظامی و امنیتی، مردم و اموال آنها، سرمایه گذاری در سیستم های امنیتی به بسیار زیاد شده است و هدف استفاده حداکثری از تمام دستاوردهای تکنولوژیکی موجود در این زمینه برای توسعه سیستم های نظارتی پیچیده است. دوربین‌های حرارتی در حال حاضر در سیستم‌های نظارت تصویری که در اماکن مختلف نصب می‌شوند، مورد استفاده قرار می‌گیرند. استفاده از دوربین های حرارتی در سیستم‌های نظارت ویدیویی بسیار مهم هستند زیرا می توانند در شرایط آب و هوایی مختلف و زمانی که زمانی که دوربین‌های معمولی RGB قابل استفاده نیستند یا نتایج ضعیفی ارائه می دهند، مورد استفاده و بهره برداری قرار گیرند. برای افزایش امنیت و نظارت در مناطق حفاظت شده شناسایی افراد و تحرکات مشکوک در شرایط آب و هوایی مختلف و در طول ۲۴ ساعت شبانه روز بسیار مهم است و توانایی تشخیص خودکار یک فرد یا شی و اعلام هشدار برای سیستم‌های نظارتی بسیار مهم است.

تاکنون، از الگوریتم‌های یادگیری ماشینی مختلفی برای شناسایی، تشخیص و ردیابی اشیایی مانند صورت انسان (کمار[1]و همکاران،۲۰۱۹) یا شکل و هیکل انسان (فرانکو[2]و همکاران،۲۰۲۰) در تصاویر نوری RGB توسعه یافته‌اند.

هدف از شناسایی و تشخیص اشیاء، دسته‌بندی و برچسب گذاری اشیا و تعیین موقعیت دقیق آنها در تصویر یا ویدئو است. امروزه مدل‌های مبتنی بر شبکه‌های عصبی کانولو شنال[3] دارای نتایج بسیار خوبی در شناسایی و تشخیص اشیاء در تصاویر RGB هستند.

با توجه به تفاوت‌ها در ویژگی‌های موجود در تصاویر بصری و حرارتی، هدف ما بررسی میزان موفقیت روش‌های یادگیری عمیق در شناسایی و تشخیص اشیا در ویدیوهای حرارتی می‌باشد.

برای شناسایی و تشخیص اشیاء در ویدیوی حرارتی از شبکه یولو۳ (ردمون[4]و فرهادی،۲۰۲۲)استفاده کرده‌ایم. برای این منظور تصاویر موجود در پایگاه داده مورد استفاده (UNIRi-TID[5]) را به فیلم تبدیل کرده و سپس اقدام به شناسایی و تشخیص افراد در ویدیوی حرارتی کرده‌ایم.

[1] Kumar
[2] Franco
[3] Convloutional Neural Network
[4] Redmon& Farhadi
[5] https://ieee-dataport.org/open-access/thermal-image-dataset-person-detection-uniri-tid

روش تحقیق

تصاویر حرارتی عمدتا برای تشخیص حضور افراد در شب یا در شرایط آب و هوایی نامساعد استفاده می شود، اما در روز که کنتراست حرارتی کافی بین افراد و محیط اطرافشان وجود ندارد، عملکرد ضعیفی دارد.

گوس و همکاران (گوس[1]و همکاران،۲۰۱۹) برای شناسایی و تشخیص عابر پیاده از تصاویر حرارتی استفاده کرده‌اند. اوسیک و همکاران(اوسیک[2]و همکاران،۲۰۱۹) آشکارساز یولو بر روی یک مجموعه داده تصویر حرارتی برای تشخیص شخص آموزش داده شده است. آن‌ها با تجزیه و تحلیل شرایط آب و هوایی مختلف، و آزمایش بر روی مجموعه‌های داده دیگر مانند حیوانات در فرآیند تست، کار خود را توسعه داده‌اند.

گمز و همکاران (گمز[3]و همکاران،۲۰۲۰)، برای شمارش افراد در فضاهای عمومی مانند کلاس درس، از تصاویر حرارتی استفاده کرده‌اند. آن‌ها ضمن ایجاد یک پایگاه داده شامل ۳۰۰۰ تصویر حرارتی، یک الگوریتم شمارش افراد بر روی تصاویر حرارتی مبتنی بر شبکه‌های عصبی کانولو شنال با قابلیت اجرا بر روی یک پلتفرم کم مصرف با حافظه محدود ایجاد کرده‌اند.

لیو و همکاران (لیو[4]و همکاران،۲۰۱۷) برای شناسایی و تشخیص اشیا در تصاویر حرارتی از شبکه‌های عصبی کانولشنی استفاده کرده‌اند. رادجر و همکاران(رادجر[5]و همکاران،۲۰۱۶) نیز با آموزش یک شبکه عصبی کانولوشنی با تصاویر حرارتی اقدام به شناسایی و تشخیص اشیاء کرده‌اند. هدف آن ها افزایش دقت تشخیص و کاهش نرخ هشدار کاذب برای یک سیستم تشخیص میان برد، با استفاده از حسگر امواج مادون قرمز است.

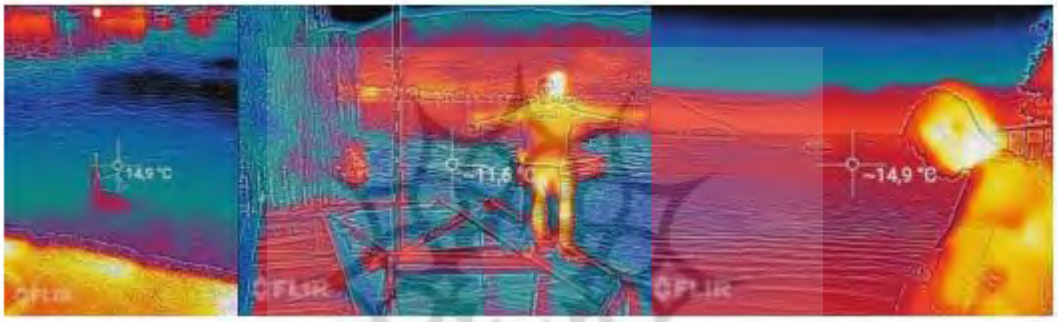
شاهید و همکارن(شاهید[6]و همکاران،۲۰۱۹) نیز برای شناسایی و تشخیص انسان در زمان واقعی در تصاویر حرارتی از شبکه‌های عصبی کانولوشنی استفاده کرده‌اند. وانگ و همکارن(وانگ[7]و همکاران،۲۰۱۹) نیز برای شناسایی و تشخیص انسان در ویدیوی حرارتی از شبکه‌های عصبی کانولوشنال عمیق استفاده کرده‌اند.

شناسایی و تشخیص انسان در ویدیو به دو دلیل ۱- منحصر به فرد بودن شکل و ضاهر هر فرد و ۲- در شب یا در شرایط آب و هوایی نامساعد چالش برانگیز است، درنتیجه در این مقاله ما شناسایی و تشخیص انان در ویدیو با استفاده از تصاویر حرارتی را مورد بررسی قرار داده‌ایم. دوربین‌های حرارتی می‌توانند گرمای تولید شده یا منعکس شده از اجسام را ثبت کرده و نتیجه را به صورت یک تصویر که بر اساس دما نقاط آن تغییر می‌کند ذخیره نمایند. آن دسته از این دوربین‌ها که در رنج میانی موج مادون قرمز و یا رنج بلند موج مادون قرمز کار می‌کنند به هیچ سورس نور خارجی

[1] Ghose
[2] Ivašić
[3] Gomez
[4] Liu
[5] Rodger
[6] Shahid
[7] Wang

دیگری برای ثبت تصاویر دارای مفهوم نیاز ندارند. به عنوان ویژگی مهم این دوربین‌ها باید توجه داشت که نتایج ضبط شده توسط آن‌ها نسبت به شرایط مختلف نوری و هوایی رباست می‌باشد. از طرف دیگر باید توجه داشت این دوربین‌ها اطلاعات بسیار کمتری را نسبت به دوربین‌های تصویربرداری عادی می‌توانند از تصویر ثبت نمایند.

برای تحلیل تصاویر دوربین‌های حراراتی لازم است توجه کنیم رنج رنگ‌های موجود در تصاویر این دوربین‌ها بر اساس دمای محیط عکس‌برداری تغییر می‌کند(شکل ۱). سردترین نقاط موجود در تصویر به صورت آبی تیره نشان داده می‌شوند و رنگ نقاط گرم‌تر تصویر به سمت سفید می‌رود.



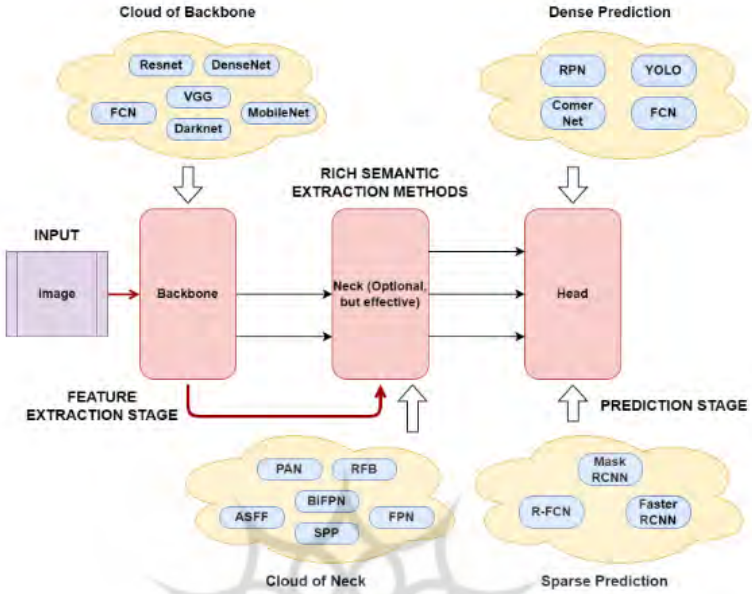
شکل ۱. نمونه ای از رنج توزیع رنگ در تصویر (کریستو[1] و همکاران،۲۰۲۰)

الگوریتم یولو

برای شناسایی و تشخیص انسان در ویدیوهای حرارتی ما در این مقاله از الگوریتم یولو استفاده کرده‌ایم. روش یولو که در سال ۲۰۱۵ ارائه شده است با استفاده از ترکیب candidate box generation و طبقه‌بندی بر اساس رگرشن توانست پروسه تشخیص اجسام را در یک مرحله انجام دهد. ساختار این روش منجر به افزایش قابل ملاحظه سرعت در تشخیص اجسام شد به نحوی که با استفاده از آن امکان پردازش تصاویر با سرعت ۴۵ فریم بر ثانیه محقق شد روش یولو۳ که در پژوهش حاضر مورد بررسی قرار گرفته است در سال ۲۰۱۸ و با اعمال تغییراتی بر روی در راستای بهبود دقت مدل‌های یولو و یولو ۲ ارائه شده است.

سه مرحله اصلی در عموم روش‌های تشخیص اجسام به صورت تک مرحله‌ای وجود دارد که این مراحل را می‌توان در شکل ۲ مشاهده نمود.

[1] Krišto



شکل ۲. ساختار کلی مدل‌های تشخیص اجسام

در گام نخست این مدل‌ها تصویر وارد یک ساختار شبکه عصبی کانولوشنی می‌شود. این ساختار که معمولا یک مدل از پیش آموزش داده شده بر روی دیتاست ImageNet می‌باشد که وظیفه استخراج ویژگی‌های تصویر را دارا می‌باشد. مدل‌های مختلفی که به منظور تسک‌های تقسیم‌بندی استفاده شده‌اند می‌توانند در انجام این وظیفه به کار گرفته شوند. در الگوریتم ورژن ۳ مدل یولو از ساختار darcknet 52 به منظور استخراج ویژگی‌ها استفاده شده است.

آموزش شبکه

در این پروژه ما ابتدا برای شبکه را با مجموعه دیتاست پاسکال ۲۰۱۲ و ۲۰۰۷ آموزش داده‌ایم. این دیتاست شامل تصاویر مربوط به اجسام مختلف در ۲۰ دسته می‌باشد. سپس به منظور شناسایی و تشخیص انسان در تصاویر حرارتی شبکه را با دیتاست UNIRi-TID مجددا آموزش می‌دهیم. بدین ترتیب علاوه بر صرفه‌جویی در زمان آموزش شبکه، مشکل کم بودن تصاویر دوربین حراراتی جهت آموزش نیز برطرف میشود. نکته نیازمند توجه در این مورد تفاوت تعداد آبجکت‌های موجود در دو دیتاست میباشد. دیتاست پاسکال شامل ۲۰ کلاس مختلف می‌باشد، در حالی که در دیتاست مورد استفاده، تنها کلاس مهم انسان می‌باشد. تفاوت در تعداد کلاس‌ها منجر به تفاوت در قسمت‌هایی از ساختار neck مدل می‌شود که مقادیر ضرایب شبکه آموزش دیده اولیه در این لایه‌ها در مدل دوم قابل استفاده نمی‌باشد. تصاویر موجود در دیتاست مورد بحث به نحوی تقسیم شده‌اند که در مجموع ۵۰۶۹ تصویر در مرحله آموزش مدل به کار گرفته شود و ۱۲۶۷ تصویر به منظور ارزیابی مدل در نظر گرفته شده است. با توجه به ابعاد در نظر گرفته شده برای تصاویر ورودی در آموزش اولیه شبکه با با مجموعه دیتاست پاسکال ۲۰۱۲ و ۲۰۰۷، سایز تصاویر موجود در دیتاست به ۴۱۶*۴۱۶ تغییر داده شده‌است. با توجه به اینکه تصاویر موجود در دیتاست در واقع فریمهای ویدیو تبدیل شده به عکس می‌باشند، به منظور پرهیز از به

کارگیری دیتای مشابه با دیتای تست در مرحله آموزش، تصاویر بر اساس شماره فریم مرتبط شده و در ادامه تصاویر ۸۰ درصد فریم‌های ابتدایی به منظور یادگیری به کار گرفته شده‌اند.

نتایج و بحث

در این بخش به ارزیابی و تحلیل روش پیشنهادی می‌پردازیم. در ابتدا مشخصات پایگاه داده تصاویر مورد استفاده را معرفی کرده و سپس روش پیشنهادی را مورد ارزیابی قرار می‌دهیم.

پایگاه‌داده تصاویر

هدف ما، شناسایی و تشخیص افراد حاضر در مناطق حفاظت شده بر اساس تصاویر حرارتی گرفته شده توسط دوربین حرارتی در شرایط هوایی مختلف(عادی، بارش سنگین و مه آلود) است. با توجه به هدف مطرح شده در گام نخست باید دیتاست مناسبی انتخاب شود که در آن تصاویر متعددی از افراد در وضعیت‌های مختلف (مانند راه رفتن، سینه‌خیز رفتن و ...) در شرایط جوی مختلف موجود باشد. به منظور دستیابی به خواسته‌های مورد بحث از دیتاست TID-UNIRi استفاده شده است. تمام سناریوهای این دیتاست در شب‌های زمستانی تصویر برداری شده‌اند. اینکار توسط دوربین‌های FLIR ThermalCam p10 WIR صورت پذیرفته است که در ارتفاع ۱۴۰ سانتیمتری از سطح زمین نصب شده است.

دید این دوربین به صورت ۱۸*۲۴ میباشد تصاویر خروجی این دوربین دارای ابعادی برابر با ۲۴۰*۳۲۰ می‌باشند که در دیتاست مورد استفاده به صورت ۹۶۰*۱۲۸۰ بازسازی شده‌اند.

معیارهای ارزیابی

ما به منظور ارزیابی عملکرد روش پیشنهادی از معیار اصلی سنجش عملکرد الگوریتم تشخیص اجسام mAP که برحسب معیارهای ارزیابی ارزیابی بازیابی۱ (رابطه ۱) و وضوح۲ (رابطه ۲) است، و معیار–اف f۳ (رابطه ۳) استفاده کرده ایم. معیار mAP یا میانگین متوسط دقت، یکی از مهمترین معیارهایی است که برای ارزیابی دقت مدل‌های شناسایی و تشخیص اشیاء استفاده می شود.

در این ارزیابی‌ها هدف شناسایی و تشخیص انسان در تصاویر حرارتی می‌باشد.

$$Recall = \frac{TP}{TP+FP} \quad (۱)$$

$$Precision = \frac{TP}{TP+FN} \quad (۲)$$

$$\mathrm{f} - measure = \frac{2*Precision*Recall}{Precision+Recall} \quad (۳)$$

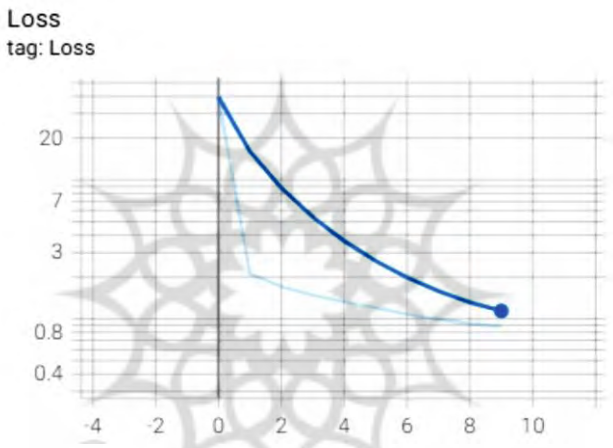
که در آن TP[۴] تعداد نمونه های مثبتی است که به درستی مثبت شناسایی شده‌اند، TN[۵] تعداد نمونه‌های منفی که به درستی منفی شناسایی شده‌اند، FP[۶] تعداد شناسایی‌های مثبت کاذب و FN[۷] تعداد شناسایی‌های منفی کاذب می‌باشند.

[1] Recall
[2] Precision
[3] F-measure
[4] True Positive
[5] True Negative
[6] False Positive
[7] False Negative

در طول پروسه یادگیری پس از طی شدن هر ۴ ایپاک بر روی دیتای آموزشی، عملکرد مدل با استفاده از دیتای تست مورد ارزیابی قرار میگیرد و در نهایت مدلی که به بیشنه مقدار mAP برسد به عنوان مدل نهایی ذخیره می‌شود. پروسه یادگیری در ۹ ایپاک صورت پذیرفته است.

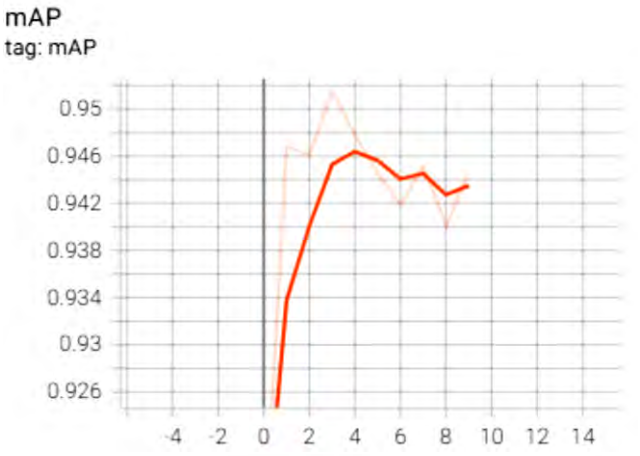
ارزیابی و مقایسه روش ارائه شده

در این بخش، نتایج ارزیابی روش پیشنهادی در شناسایی و تشخیص انسان در ویدیوهای حرارتی ارائه شده است. با توجه به شکل ۳، با شروع فرآیند یادگیری تابع هزینه با بصورت پیوسته روند همگرایی به سمت مقادیر کمتر را طی کرده است. این بدان معناست که الگوریتم بهینه‌سازی تعریف شده در کمینه کردن مقدار خطا موفق عمل کرده است.



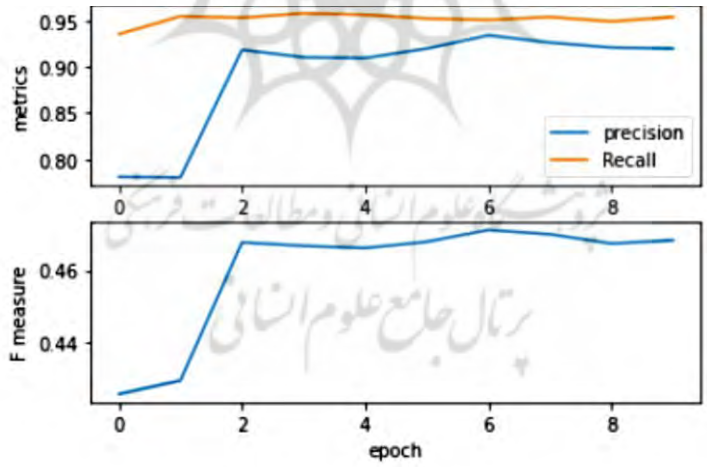
شکل۳ . مقدار تابع هزینه دیتای یادگیری بر حسب ایپاک

با وجود روند کاهشی تابع هزینه با توجه به نمودار mAP (شکل ۴) نشان می‌دهد که پس از ایپاک چهارم مقدار دقت مدل روند منفی را آغاز می‌کند که این مو ضوع نا شی از اورفیت شدن مدل بر روی دیتای یادگیری می‌با شد. به همین دلیل در این پیاده-سازی تعداد ایپاک کمتری مورد استفاده قرار گرفته است.



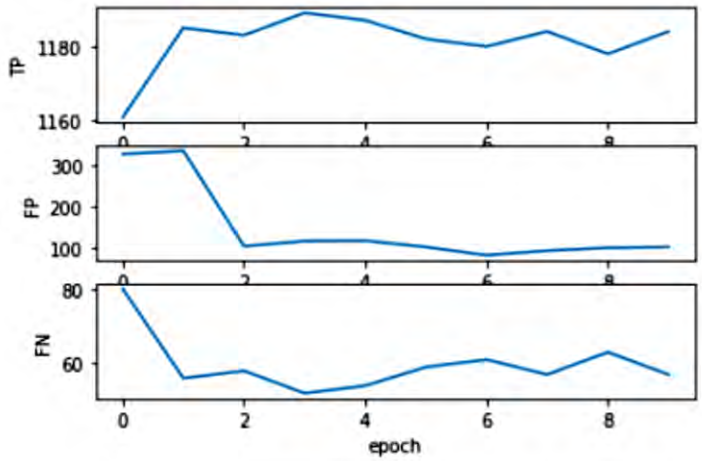
شکل ٤. نمودار تغییرات map بر حسب ایپاک روی تصاویر حرارتی (نارنجی: نمودار smooth شده تغییرات، سایه: نمودار حقیقی تغییرات map)

نتایج حا صل از ارزیابی روش پی شنهادی روی ویدیوی حرارتی برح سب معیارهای بازیابی و و ضوح نیز در شکل ۵ ارائه شده‌است.



شکل ۵. ارزیابی نتایج حاصل بر روی تصاویر حرارتی

در شکل ۶ نیز مقادیر مثبت حقیقی، مثبت کاذب و منفی کاذب برح سب ایپاک، با ا ستفاده از مجموعه داده ت ست ارائه شده‌است. عملکرد مدل به مرور و بر حسب ایپاک‌های طی شده در تشخیص مقادیر مثبت کاذب عملکرد بسیار بهتری را از خود نشـــان می‌دهد. هر چند از ایپاک ۴ به بعد عملکرد مدل در تشـــخیص مقادیر مثبت حقیقی روند منفی را شـــروع می‌کند، با این حال تغییر این عدد به نسبت مقادیر مربوط به منفی کاذب قابل توجه نیست.



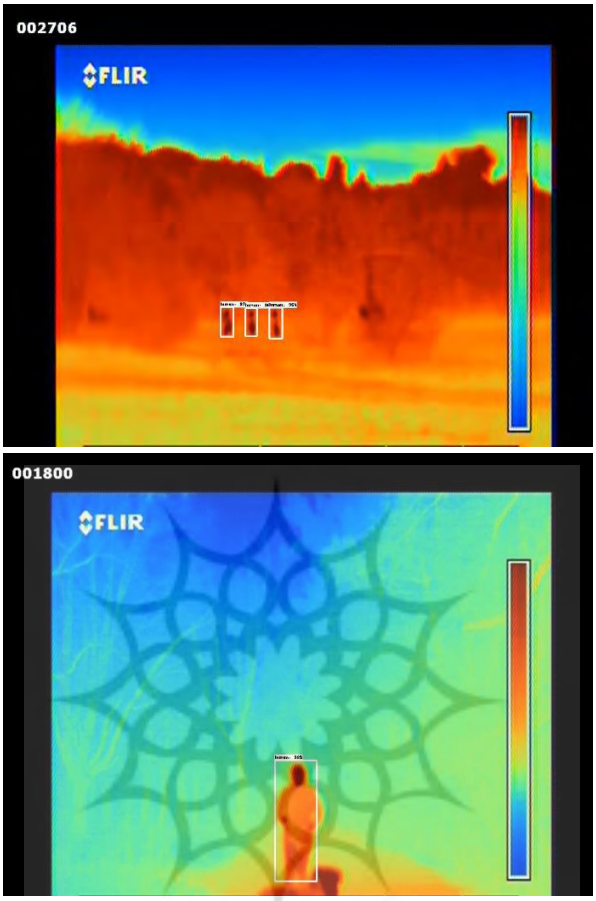
شکل ٦. مقادیر مثبت حقیقی، مثبت کاذب و منفی کاذب نسبت به ایپاک

در جدول شماره ۱ نیز نتیجه مقایسه روش پیشنهادی با کار انجام شده قبلی ارائه شده است. براساس نتایج ارائه شده روش پیشنهادی در مقایسه با کار انجام شده قبلی از دقت بهتری برخوردار است.

جدول ۱. مقایسه نتایج روش پیشنهادی با کار انجام شده قبلی در شناسایی انسان در تصاویر حرارتی.

روش	AP
SSD(W. Liu et al., 2016)	۹۴,۰۲ درصد
روش پیشنهادی	۹۵,۴۸ درصد

و نهایتا در شکل شماره ۷ تصاویر شناسایی و تشخیص انسان در ویدیوی دوربین حرارتی با استفاده از مدل آموزش داده شده ارائه شده است.



شکل ۷. نمونه‌ای از تشخیص انسان توسط دوربین حرارتی

نتیجه گیری

به منظور جلوگیری از حوادث ناخواسته و محافظت از اماکن خصوصی و امنیتی، مردم و اموال آنها، سرمایه گذاری در سیستم های نظارتی و امنیتی بسیار زیاد شده است و هدف استفاده حداکثری از تمام دستاوردهای تکنولوژیکی موجود در این زمینه برای توسعه سیستم های نظارتی پیچیده است.

با پیشرفت فناوری و استفاده گسترده از تصاویر حرارتی، شاهد روند تحولی در حوزه تشخیص و شناسایی اشیا هستیم. استفاده از دوربین‌های حرارتی امکان شناسایی و تشخیص انسان و اشیاء مختلف را در شرایط آب و هوایی مختلف و یا در طول شب برای ما فراهم می‌کند.

ما در این مقاله از شبکه یولو۳ برای شناسایی و تشخیص اشیاء در استفاده کرده‌ایم و براساس نتایج حاصل با mAP = 94.6، شناسایی و تشخیص انسان را انجام داده‌ایم.

منابع

Franco, A., Magnani, A., & Maio, D. (2020). A multimodal approach for human activity recognition based on skeleton and RGB data. *Pattern Recognition Letters*, *131*, 293–299.

Ghose, D., Desai, S. M., Bhattacharya, S., Chakraborty, D., Fiterau, M., & Rahman, T. (2019). Pedestrian detection in thermal images using saliency maps. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 0–0. http://openaccess.thecvf.com/content_CVPRW_2019/html/PBVS/Ghose_Pedestrian_Detection_in_Thermal_Images_Using_Saliency_Maps_CVPRW_2019_paper.html

Gomez, A., Conti, F., & Benini, L. (2018). Thermal image-based CNN's for ultra-low power people recognition. *Proceedings of the 15th ACM International Conference on Computing Frontiers*, 326–331. https://doi.org/10.1145/3203217.3204465

Ivašić-Kos, M., Krišto, M., & Pobar, M. (2019). Human Detection in Thermal Imaging Using YOLO. *Proceedings of the 2019 5th International Conference on Computer and Technology Applications*, 20–24. https://doi.org/10.1145/3323933.3324076

Krišto, M., Ivasic-Kos, M., & Pobar, M. (2020). Thermal object detection in difficult weather conditions using YOLO. *IEEE Access*, *8*, 125459–125476.

Kumar, A., Kaur, A., & Kumar, M. (2019). Face detection techniques: A review. *Artificial Intelligence Review*, *52*(2), 927–948. https://doi.org/10.1007/s10462-018-9650-2

Liu, Q., Lu, X., He, Z., Zhang, C., & Chen, W.-S. (2017). Deep convolutional neural networks for thermal infrared object tracking. *Knowledge-Based Systems*, *134*, 189–198.

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., & Berg, A. C. (2016). SSD: Single Shot MultiBox Detector. In B. Leibe, J. Matas, N. Sebe, & M. Welling (Eds.), *Computer Vision – ECCV 2016* (Vol. 9905, pp. 21–37). Springer International Publishing. https://doi.org/10.1007/978-3-319-46448-0_2

Redmon, J., & Farhadi, A. (2022). YOLOv3: An Incremental Improvement, ArXiv. org.(2018). *arXiv Preprint arXiv:1804.02767*.

Rodger, I., Connor, B., & Robertson, N. M. (2016). Classifying objects in LWIR imagery via CNNs. *Electro-Optical and Infrared Systems: Technology and Applications XIII*, *9987*, 152–165. https://www.spiedigitallibrary.org/conference-proceedings-of-spie/9987/99870H/Classifying-objects-in-LWIR-imagery-via-CNNs/10.1117/12.2241858.short

Shahid, N., Yu, G.-H., Trinh, T. D., Sin, D.-S., & Kim, J.-Y. (2019). Real-time implementation of human detection in thermal imagery based on CNN. 한국정보기술학회논문지, *17*(1), 107–121.

Wang, X., & Hosseinyalamdary, S. (2019). Human detection based on a sequence of thermal images using deep learning. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *42*, 127–132.