

Using the YOLOv12 Model for Verifying the Correct Color Sequence of Wires in Network Cables (Patch Cords) on the Production Line

Amin Doroodchi*

Computer Department, Islamic Azad University,
Beyza branch, Beyza, Iran.

Danial Soleymani

R&D at Nedaye Sabz Company, Isfahan Branch,
Iran.

Abstract

In the production process of network cables, ensuring the correct color sequence of wire pairs inside the standard connector plays a critical role in the final performance of the cable, as any misplacement or color-ordering error can lead to defective products and impose significant costs. Traditional inspection methods based on visual examination through digital microscopes are typically time-consuming, tedious, and prone to human error. In this study, an intelligent system based on the twelfth version of the YOLO object detection model was developed to identify the position and verify the correct color sequence of wires in patch cords. The dataset used consisted of 2,500 images captured from microscopic views of network connectors, which were divided into 70% for training, 15% for validation, and 15% for testing. The proposed model, leveraging a single-stage architecture and attention mechanisms during learning, achieved highly accurate wire detection with approximately 98% precision. Additionally, the overall mean accuracy, classification precision, and recall were around 95%, 99%, and 98%, respectively. The results demonstrate that this system can reliably and in real time verify the correctness of wire color sequencing on the production line without the need for human intervention, thereby reducing human error and enhancing efficiency in the manufacturing process.

Keywords: automated inspection, wire color-sequence detection, deep learning, computer vision, YOLO, network cable, production quality control

استفاده از مدل YOLOv12 برای تشخیص صحت ترتیب رنگ سیم‌ها در کابل‌های شبکه (پچ کورد) در خط تولید

امین درودچی*

گروه کامپیوتر، دانشگاه آزاد اسلامی واحد بیضا، بیضا، ایران.

دانیال سلیمانی

واحد تحقیق و توسعه، شرکت ندای سبز، اصفهان، ایران.

چکیده

در فرایند تولید کابل‌های شبکه، اطمینان از صحت ترتیب رنگ رشته‌سیم‌ها در داخل کانکتور استاندارد نقش بسیار مهمی در عملکرد نهایی کابل دارد؛ چراکه هرگونه جابه‌جایی یا اشتباه در رنگ‌بندی می‌تواند منجر به تولید محصول معیوب و تحمیل هزینه‌های قابل توجه شود. روش‌های سنتی بازرسی که مبتنی بر مشاهده‌ی چشمی از طریق میکروسکوپ دیجیتال هستند، معمولاً زمان‌بر، خسته‌کننده و مستعد خطاهای انسانی هستند. در این پژوهش، یک سامانه‌ی هوشمند بر پایه‌ی مدل شناسایی اشیای نسخه دوازدهم یولو^۱ به منظور تشخیص موقعیت و ترتیب رنگ صحیح رشته‌سیم‌ها در کابل‌های پچ کورد توسعه داده شده است. مجموعه داده مورد استفاده شامل دو هزار و پانصد تصویر تهیه‌شده از نمای میکروسکوپی کانکتورهای شبکه بود که به نسبت هفتاد درصد برای آموزش، پانزده درصد برای اعتبارسنجی و پانزده درصد برای آزمون تقسیم شدند. مدل پیشنهادی با بهره‌گیری از ساختار تک‌مرحله‌ای و مکانیزم توجه در مرحله‌ی یادگیری، موفق به شناسایی دقیق سیم‌ها با دقت حدود نود و هشت درصد گردید. همچنین، میانگین دقت کلی، دقت طبقه‌بندی و نرخ بازیابی به ترتیب در حدود نود و پنج درصد، نود و نه درصد و نود و هشت درصد به دست آمدند. نتایج نشان می‌دهند که این سامانه می‌تواند به صورت بلادرنگ، قابل اعتماد و بدون نیاز به مداخله‌ی انسانی، درستی ترتیب رنگ سیم‌ها را در خط تولید بررسی کرده و موجب کاهش خطاهای انسانی و افزایش بهره‌وری در فرایند تولید گردد.

کلیدواژه‌ها: بازرسی خودکار، تشخیص ترتیب رنگ سیم، یادگیری عمیق، بینایی ماشین، یولو، کابل شبکه، کنترل کیفیت تولید

مقدمه

در فرایند تولید صنعتی کابل‌های شبکه، صحت چیدمان رشته‌سیم‌ها در داخل کانکتور از اهمیت حیاتی برخوردار است؛ چراکه هرگونه جابه‌جایی در ترتیب رنگ این سیم‌ها منجر به تولید کابل‌های معیوب و ناکارآمد خواهد شد. در استانداردهای سیم‌بندی، نظیر T568A و T568B، ترتیب هشت رشته‌سیمی که از طریق کانکتور به شبکه متصل می‌شوند، به صورت دقیق تعریف شده و هرگونه خطا در این مرحله ممکن است منجر به اختلال کامل در عملکرد شبکه شود (Tian et al., 2025). این در حالی است که در بسیاری از خطوط تولید، بررسی صحت این ترتیب رنگ‌ها همچنان به صورت دستی و توسط نیروی انسانی انجام می‌شود، به گونه‌ای که اپراتورها با کمک بزرگ‌نمایی نوری (اغلب از طریق میکروسکوپ دیجیتال)، رنگ رشته‌ها را به صورت چشمی ارزیابی می‌کنند. این روش نه تنها زمان‌بر، بلکه تحت تأثیر خستگی اپراتور، نور محیط، مهارت فردی و محدودیت دقت بینایی انسان بوده و در نتیجه مستعد بروز خطا می‌باشد (Wang et al., 2024).

به منظور کاهش خطاهای انسانی، افزایش سرعت و دستیابی به کنترل کیفیت یکنواخت، استفاده از سامانه‌های بازرسی خودکار نوری^۱ در بسیاری از حوزه‌های تولید صنعتی توسعه یافته است. البته، لزوم به کارگیری سیستم‌های کنترلی و هوشمند به جای نظارت دستی، تنها محدود به صنعت کابل‌سازی نیست، بلکه در سایر کاربردهای حساس مانند تجهیزات پزشکی نیز برای پیشگیری از آسیب و مدیریت دقیق فشار، از رویکردهای مشابهی استفاده شده است که اهمیت توسعه‌ی چنین سامانه‌هایی را دوچندان می‌کند (Shayan et al., 2021). در این میان، ترکیب روش‌های بینایی ماشین^۲ با یادگیری عمیق^۳ به راهکاری مؤثر و قابل اعتماد در حوزه‌های مختلف تبدیل شده است (Yuan et al., 2024; Zhao et al., 2025). در حوزه‌ی تشخیص و ارزیابی ترتیب سیم‌های رنگی نیز به کارگیری شبکه‌های عصبی کانولوشنی^۴ امکان تحلیل دقیق تصاویر، استخراج ویژگی‌های رنگی و موقعیت مکانی سیم‌ها را فراهم می‌سازد (Liu et al., 2025). اگرچه توسعه‌ی این سامانه‌ها نیازمند حجم مناسبی از داده‌های آموزشی و قدرت پردازشی بالا برای آموزش مدل‌ها است، اما در صورت طراحی صحیح می‌توان آن‌ها را در خطوط تولید واقعی و به صورت بلادرنگ^۵ پیاده‌سازی کرد (Wang et al., 2024).

در میان الگوریتم‌های یادگیری عمیق، خانواده‌ی مدل‌های یولو یکی از پرکاربردترین و سریع‌ترین روش‌ها برای تشخیص آنی اشیاء در تصویر محسوب می‌شود (Redmon et al., 2016). مدل‌های یولو به دلیل معماری تک‌مرحله‌ای خود، در مقایسه با روش‌های دقیق‌تر اما کندتر، توازن بسیار مطلوبی بین سرعت و دقت فراهم می‌آورند. این ویژگی موجب شده است که یولو به گزینه‌ای جذاب برای کاربردهای صنعتی از جمله بازرسی خودکار در خطوط تولید تبدیل شود (Li et al., 2025). در نسخه‌های اخیر این مدل، به ویژه نسخه دوازدهم (یولو ۱۲)، با استفاده از مکانیزم‌های توجه^۶، ماژول‌های تجمیع ویژگی‌های چندمقیاسی و بهینه‌سازی توابع اتلاف^۷، دقت تشخیص حتی برای اشیاء کوچک یا مشابه به طور چشمگیری افزایش یافته است (Tian et al., 2025). این نسخه از یولو ضمن حفظ سرعت پردازش بسیار بالا (در حد چند میلی‌ثانیه)، قادر است دقتی در سطح مدل‌های پیچیده‌تر و سنگین‌تر ارائه دهد؛ از این رو، گزینه‌ای مناسب برای کاربردهای صنعتی با حجم تولید بالا به شمار می‌رود (Mao & Hong, 2025; Nguyen et al., 2024).

1. automatic optical inspection
2. machine vision
3. deep learning
4. convolutional neural networks
5. real-time
6. attention mechanisms
7. loss functions

با توجه به اهمیت تشخیص صحیح رنگ و ترتیب سیم‌ها در کیفیت نهایی کابل شبکه و محدودیت‌های روش‌های سنتی بازرسی چشمی، در این پژوهش یک سامانه‌ی بینایی ماشین مبتنی بر مدل یولو نسخه دوازدهم طراحی و پیاده‌سازی شده است. همچنین، رویکردهای یادگیری چندحالتی^۱ که داده‌های دیداری را با داده‌های متنی یا دیگر حسگرها ترکیب می‌کنند نیز در حال تبدیل شدن به یک استاندارد برای افزایش دقت در حوزه‌های حساسی مانند تشخیص پزشکی هستند (Alaei et al., 2025). هدف از این پژوهش آن است که بتوان از طریق آموزش مدل بر پایه‌ی داده‌های واقعی تولیدشده در خط کابل‌سازی، یک الگوریتم دقیق، سریع و خودکار برای ارزیابی چیدمان رنگ سیم‌ها در کانتورهای RJ45 توسعه داد. این سامانه پس از شناسایی سیم‌ها در تصویر، ترتیب آن‌ها را از چپ به راست مقایسه کرده و صحت انطباق آن با الگوی استاندارد را تعیین می‌کند. در ادامه، ضمن مرور پژوهش‌های مرتبط، روش کار، نحوه‌ی آماده‌سازی داده‌ها، ساختار مدل، نتایج ارزیابی و تحلیل عملکرد ارائه خواهد شد.

مرور پژوهش‌های پیشین

تشخیص خودکار ترتیب رنگ سیم‌ها در کابل‌های چندرشته‌ای یکی از چالش‌های فنی دقیق در حوزه‌ی کنترل کیفیت صنعتی است که به صورت تخصصی کمتر در مطالعات مستقل بررسی شده و اغلب در قالب بخشی از سامانه‌های بازرسی کابل یا سیم‌بندها مورد توجه قرار گرفته است. یکی از نخستین تلاش‌ها در این زمینه را تسای و چنگ^۲ (۲۰۲۲) انجام داده‌اند. آن‌ها یک سامانه‌ی بازرسی خودکار نوری مبتنی بر پردازش تصویر کلاسیک برای تشخیص توالی رنگ رشته‌سیم‌ها ارائه کردند که با تحلیل لبه‌ها و گرادیان‌های افقی در تصاویر و سپس استفاده از منطق تصمیم‌گیری ترتیبی، توانستند محل و ترتیب رنگ سیم‌ها را استخراج کنند. این روش، اگرچه در شرایط کنترل‌شده دقتی در حدود ۹۸/۵ درصد داشت، اما به شدت به کیفیت نور و وضوح تصویر وابسته بود؛ به طوری که با کاهش طول نمای قابل رؤیت سیم‌ها یا تغییر رنگ پس‌زمینه، دقت تا کمتر از ۹۳ درصد افت می‌کرد. همچنین، سرعت پردازش پایین (حدود ۱/۷ ثانیه برای هر تصویر) که آن را برای خطوط تولید سریع غیر کاربردی می‌کرد (Tsai & Cheng, 2022).

در مقابل، روش‌های مبتنی بر یادگیری عمیق و به‌ویژه شبکه‌های عصبی کانولوشنی توانسته‌اند بر بسیاری از محدودیت‌های روش‌های سنتی غلبه کنند. همچنین، کاربرد شبکه‌های عصبی عمیق در تشخیص الگوهای مکانی و تخمین موقعیت، حتی در فضای سه‌بعدی و برای ژست‌های پیچیده، نشان‌دهنده‌ی توانایی بالای این مدل‌ها در استخراج اطلاعات دقیق مکانی است (Mirzadeh & Zare, 2023). این مدل‌ها با یادگیری ویژگی‌های مکانی و رنگی سیم‌ها، قدرت تعمیم بالاتری با شرایط متنوع دارند. برای مثال، ژائو و همکاران^۳ (۲۰۲۴) از ترکیب شبکه‌های کانولوشنی و لایه‌های توجه برای تشخیص ترتیب سیم‌ها در کابل‌های برق استفاده کردند و به دقتی فراتر از ۹۶ درصد با سرعت پردازش بالاتر از بلادرنگ دست یافتند. در پژوهشی دیگر، یوان و همکاران^۴ (۲۰۲۵) سامانه‌ای مبتنی بر یولو نسخه هشتم ارائه کردند که برای تشخیص اتصال سیم به ترمینال‌های الکتریکی طراحی شده بود و توانست وجود یا عدم وجود سیم در موقعیت موردنظر را با دقت بالای ۹۹ درصد تشخیص دهد.

علاوه بر کاربردهای مشابه در سیم‌بندها، در بازرسی سطوح سیم‌ها و کابل‌های صنعتی نیز مدل‌های یولو با اصلاحات مختلف به‌طور گسترده استفاده شده‌اند. در پژوهش مائو و هونگ^۵ (۲۰۲۵) کاربرد مدل یولو در شناسایی عیوب بافت پارچه و سیم‌های صنعتی بررسی شد. آن‌ها با استفاده از نسخه‌های مختلف یولو (از نسخه اول تا یازدهم)

1. multimodal learning
2. Tsai & Cheng
3. Zhao et al.
4. Yuan et al.
5. Mao & Hong

به این نتیجه رسیدند که دقت تشخیص نواقص، با ارتقای نسخه مدل و افزودن مکانیزم‌های توجه، به‌طور پیوسته بهبود یافته است.

از منظر فنی، تکامل مدل یولو از نسخه اول تا دوازدهم نشان‌دهنده‌ی رشد قابل توجهی در توانمندی‌های آن برای کاربردهای بلادرنگ و دقیق است. نسخه دوم یولو با معرفی جعبه‌های لنگر^۱ و تعیین خودکار ابعاد آن‌ها، دقت مکان‌یابی را نسبت به نسخه اول بهبود داد (Redmon & Farhadi, 2017). نسخه‌های سوم و چهارم با بهره‌گیری از معماری شبکه‌های چندمقیاسی مانند PANet و تکنیک‌های داده‌افزایی نوآورانه نظیر Mosaic، به دقتی بالاتر در شناسایی اشیای کوچک و مشابه دست یافتند (Bochkovskiy et al., 2020). در ادامه، نسخه‌های پنجم تا هشتم (YOLOv5 تا YOLOv8) نیز بر بهینه‌سازی سرعت، کاهش حجم مدل و حذف وابستگی به جعبه‌های لنگر متمرکز شدند و ساختارهایی بدون لنگر^۲ را توسعه دادند که هم انعطاف‌پذیرتر و هم ساده‌تر برای آموزش بودند (Jocher et al., 2023).

در تازه‌ترین تحولات، تیان و همکاران^۳ (۲۰۲۵) مدل یولو نسخه دوازدهم را با تمرکز بر تلفیق بینایی کانولوشنی و مکانیزم‌های توجه ارائه کردند. این مدل با افزودن ماژول‌های توجه ناحیه‌ای (A2)، ساختارهای تجمیعی نظیر Residual-ELAN و بهینه‌سازی‌های حافظه مانند FlashAttention، هم دقت بالاتری نسبت به نسخه‌های پیشین و هم زمان استنتاج بسیار پایین (در حد ۱ تا ۲ میلی‌ثانیه) فراهم می‌سازد. به‌عنوان نمونه، نسخه YOLOv12-Nano بر روی مجموعه داده استاندارد COCO موفق به دستیابی به دقت میانگین بالاتر از ۴۰ درصد شده است (Nguyen et al., 2024).

پژوهش لی و همکاران^۴ (۲۰۲۵) نشان داد که نسخه دوازدهم یولو حتی در مقیاس صنعتی نیز توانایی پیاده‌سازی دارد و در آزمایش‌های میدانی، دقت ۹۷/۸ درصدی در تشخیص الگوهای رنگی پیچیده و سرعت پردازش بالاتر از ۱۲۰ فریم بر ثانیه را ارائه داده است. در نتیجه، ترکیب این نسخه از یولو با داده‌های واقعی می‌تواند راهکاری مناسب برای بازرسی خودکار ترتیب رنگ سیم‌ها در خط تولید کابل‌های شبکه باشد.

روش کار

در این بخش قابل ذکر است که روش کار این پژوهش، دیتاست. لذا، مدل انتخاب شده به همراه محیط کاری انتخاب شده بررسی خواهد شد.

داده‌های آموزشی و پیش‌پردازش

برای آموزش و ارزیابی مدل پیشنهادی، یک مجموعه داده اختصاصی از تصاویر سر کابل‌های شبکه استاندارد تهیه گردید. بدین منظور تعداد دو هزار و پانصد تصویر رنگی با وضوح تقریبی ۱۲۰۰ در ۱۶۰۰ پیکسل با استفاده از میکروسکوپ دیجیتال صنعتی از نمای روبه‌روی کانکتورهای شبکه (RJ45) ثبت شد. هر تصویر شامل نمایی واضح از هشت رشته سیم رنگی قرار گرفته در درون سوکت شفاف RJ45 بود که مطابق با استاندارد T568B، رنگ‌های سفید-نارنجی، نارنجی، سفید-سبز، آبی، سفید-آبی، سبز، سفید-قهوه‌ای و قهوه‌ای را شامل می‌شدند.

1. anchor boxes
2. anchor-free
3. Tian et al.
4. Li et al.

در مرحله‌ی بعد، کلیه‌ی تصاویر به‌صورت دستی نشانه‌گذاری شدند. به این شکل که پیرامون هر رشته‌سیم یک جعبه محصورکننده^۱ رسم گردید و برچسب مربوط به رنگ آن ثبت شد. در مجموع، هشت کلاس رنگی تعریف گردید و داده‌ها با نسبت هفتاد درصد برای آموزش، پانزده درصد برای اعتبارسنجی و پانزده درصد برای آزمون تقسیم شدند.

برای افزایش تعمیم‌پذیری مدل و جلوگیری از بیش‌برازش^۲، از روش‌های افزون‌سازی داده^۳ بهره گرفته شد. این عملیات شامل چرخش جزئی، برش و بزرگ‌نمایی محدود، معکوس‌سازی افقی، تغییر روشنایی و تنظیم کنتراست تصاویر بود. این اقدامات سبب می‌شود که مدل بتواند به‌خوبی با تنوعات ظاهری در شرایط واقعی خط تولید سازگار شود. در شکل ۱ نمونه‌ای از تصویربرداری از کابل‌ها را مشاهده می‌کنید.



شکل ۱. نمونه تصویر گرفته شده از کابل‌ها و ترتیب آن‌ها در شیراز آزمایشگاهی

معماری مدل و تنظیمات آموزش

مدل انتخاب شده برای این پژوهش، نسخه دوازدهم مدل یولو ۱۲ بود که به‌صورت متن‌باز در بستر پای‌تورچ^۴ پیاده‌سازی شده است. یولو نسخه دوازدهم از معماری تک‌مرحله‌ای را برای تشخیص اشیاء بهره می‌برد و مستقیماً مختصات جعبه‌های محصورکننده و کلاس‌های مربوط به اشیاء را در یک مرحله خروجی می‌دهد.

ساختار این مدل شامل سه بخش اصلی بود:

۱. استخراج ویژگی‌ها^۵ که در آن از مکانیزم توجه ناحیه‌ای^۶ برای تمرکز بر بخش‌های مهم تصویر استفاده شد؛
۲. تجمیع ویژگی‌ها با استفاده از ساختار باقیمانده ELAN^۷ جهت حفظ اطلاعات در عمق شبکه و بهبود گرادیان در فرایند آموزش؛

1. bounding box
 2. overfitting
 3. data augmentation
 4. pytorch
 5. backbone
 6. region-based attention
 7. residual-ELAN

۳. پیش‌بینی چندمقیاسی که در سه سطح از خروجی شبکه، اشیا با اندازه‌های مختلف از جمله سیم‌های نازک شناسایی می‌شدند.

برای تطبیق مدل با کاربرد موردنظر، خروجی مدل به گونه‌ای پیکربندی شد که بتواند هشت کلاس رنگی سیم را شناسایی کند. اگرچه اندازه‌ی سیم‌ها در تصاویر تقریباً یکسان بود، اما فعال‌سازی سه مقیاس خروجی موجب بهبود دقت در شرایط مختلف شد. همچنین، از جعبه‌های لنگر^۱ استفاده گردید که با الگوریتم خوشه‌بندی k-میانگین^۲ بر اساس ابعاد واقعی سیم‌ها بهینه‌سازی شدند. در نهایت، نه جعبه لنگر در سه سطح شبکه توزیع گردید تا موقعیت‌یابی دقیق‌تری حاصل شود.

تابع اتلاف مدل شامل سه مؤلفه بود:

۱. اتلاف مکان‌یابی^۳ با استفاده از تابع هم‌پوشانی کامل CIoU برای کاهش خطا در مختصات پیش‌بینی شده؛

$$L_{CIoU} = 1 - IoU + \frac{\rho^2(b, b^{gt})}{c^2} + \alpha v \quad (1)$$

در این پژوهش، از تابع اتلاف CIoU^۴ برای کاهش خطای مکان‌یابی استفاده شد. این تابع که در فرمول (۱) نمایش داده شده است، علاوه بر اندازه‌ی هم‌پوشانی (IoU) بین جعبه پیش‌بینی شده (b) و جعبه حقیقت زمین (b^{gt})، به فاصله‌ی مراکز آن‌ها ($\frac{\rho^2(b, b^{gt})}{c^2}$) و همچنین به میزان ثابت در نسبت ابعادی (جمله αv) توجه می‌کند. این رویکرد، در مقایسه با روش‌های سنتی‌تر، منجر به همگرایی سریع‌تر و دقیق‌تر در پیش‌بینی مختصات می‌شود.

۲. اتلاف تشخیص شی^۵ بر اساس باینری کراس‌انترپی برای تعیین وجود یا عدم وجود شی در هر موقعیت پیشنهادی؛

۳. اتلاف طبقه‌بندی^۶ برای تعیین رنگ صحیح که با تکنیک نرم‌سازی برچسب^۷ بهبود یافته بود. موفقیت در طبقه‌بندی نیازمند ارتباط دقیق بین ویژگی‌های بصری استخراج شده توسط مدل و برچسب‌های معنایی (رنگ) است؛ موضوعی که در تحقیقات جدید بر روی نمایش معنایی تصاویر نیز مورد بررسی قرار گرفته است (Zare, 2025).

ترکیب وزنی این مؤلفه‌ها به گونه‌ای تنظیم شده است که حدود پنجاه درصد برای مکان‌یابی، چهل درصد برای تشخیص شی و ده درصد برای طبقه‌بندی اختصاص یابد.

$$L_{Total} = \lambda_{Loc}L_{CIoU} + \lambda_{Obj}L_{Objectness} + \lambda_{Cls}L_{Classification} \quad (2)$$

تابع اتلاف نهایی مدل (L_{Total}) که در فرآیند بهینه‌سازی کمینه می‌گردد، ترکیب وزنی از سه مؤلفه‌ی اصلی است که در فرمول (۲) نشان داده شده است. ضریب‌های وزنی (λ) تعیین‌کننده‌ی اهمیت هر بخش بودند. بر این اساس، بیشترین تمرکز مدل با ضریب $\lambda_{Loc} = 0.5$ بر صحت مکان‌یابی، سپس با $\lambda_{Obj} = 0.4$ بر توانایی تشخیص وجود شی (رشته‌سیم) و در نهایت با $\lambda_{Cls} = 0.1$ بر صحت طبقه‌بندی رنگ سیم‌ها قرار داده شد. این توزیع وزن، اولویت‌بندی پژوهش را بر دقت موقعیت‌یابی و تشخیص نشان می‌دهد.

1. anchor boxes
2. k-means clustering
3. localization loss
4. complete intersection over union
5. objectness loss
6. classification loss
7. label smoothing

معیار اصلی ارزیابی مدل در طول آموزش، میانگین دقت متوسط^۱ در مجموعه‌ی اعتبارسنجی در سطوح مختلف هم‌پوشانی^۲ بود.

$$mAP = \frac{1}{C} \sum_{i=1}^C AP_i \quad (3)$$

معیار اصلی ارزیابی عملکرد مدل، میانگین دقت متوسط (mAP) در سطح آستانه‌ی هم‌پوشانی مشخص (mAP@IoU) بود که در فرمول (۳) تعریف می‌شود. C تعداد کل کلاس‌های رنگی (برابر ۸) است و AP_i نشان‌دهنده‌ی دقت متوسط (Average Precision) برای کلاس i-ام است. AP_i در واقع مساحت زیر نمودار دقت-فراخوانی است. استفاده از mAP تضمین می‌کند که مدل هم در شناسایی صحیح (دقت بالا) و هم در یافتن تمامی موارد مرتبط (فراخوانی بالا) برای هر هشت رنگ سیم عملکرد مطلوبی داشته باشد.

فرایند آموزش

آموزش مدل در محیط ابری گوگل کولب^۳ با استفاده از واحد پردازش گرافیکی تسلا T4^۴ انجام شد. وزن‌های اولیه از مدل از پیش آموزش دیده شده‌ی یولو بر روی پایگاه داده عمومی کوکو^۵ بارگذاری شدند تا آموزش از نقطه‌ای بهینه آغاز شود.

نرخ اولیه یادگیری برابر با ۰/۰۰۱ در نظر گرفته شد که در طی پنجاه دوره آموزشی^۶ به صورت پله‌ای کاهش یافت. اندازه‌ی دسته‌های داده^۷ به دلیل محدودیت حافظه گرافیکی، برابر ۱۶ انتخاب شد. در پایان هر دوره، عملکرد مدل روی داده‌های اعتبارسنجی ثبت گردید و نمودارهای یادگیری شامل روند کاهش تابع اتلاف و افزایش دقت ترسیم شدند. همچنین، برای جلوگیری از بیش‌برازش، در صورت عدم بهبود mAP طی پنج دوره متوالی، آموزش به صورت خودکار متوقف می‌شد (توقف زودهنگام)^۸. زمان آموزش هر دوره حدود یک تا دو دقیقه و مجموع زمان آموزش کامل حدود یک ساعت برآورد گردید.

الگوریتم تصمیم‌گیری و تطبیق با الگوی استاندارد

پس از پایان آموزش، برای استفاده عملی مدل در خط تولید، یک الگوریتم ساده اما دقیق برای ارزیابی ترتیب رنگ سیم‌ها طراحی شد. در این الگوریتم، ابتدا خروجی مدل شامل مختصات مکانی و رنگ هر رشته سیم استخراج می‌شود. سپس با مرتب‌سازی سیم‌ها بر اساس مؤلفه افقی موقعیت آن‌ها (محور x)، توالی رنگ‌ها به دست می‌آید.

در گام بعد، توالی پیش‌بینی شده با الگوی استاندارد موردنظر (نظیر T568B) مقایسه می‌شود. اگر تطابق کامل وجود داشته باشد، کابل سالم و دارای سیم‌بندی صحیح شناخته می‌شود؛ در غیر این صورت، کابل به عنوان معیوب برچسب گذاری می‌شود و برای بررسی بیشتر کنار گذاشته خواهد شد. این فرایند به دلیل سادگی و سرعت بالا، هیچ بار محاسباتی قابل توجهی نسبت به مرحله‌ی شناسایی ندارد و قابلیت اجرا به صورت بلادرنگ در خط تولید را داراست.

-
1. Mean average precision (mAP)
 2. IoU
 3. google colab
 4. nvidia tesla T4
 5. COCO
 6. epoch
 7. batch size
 8. early stopping

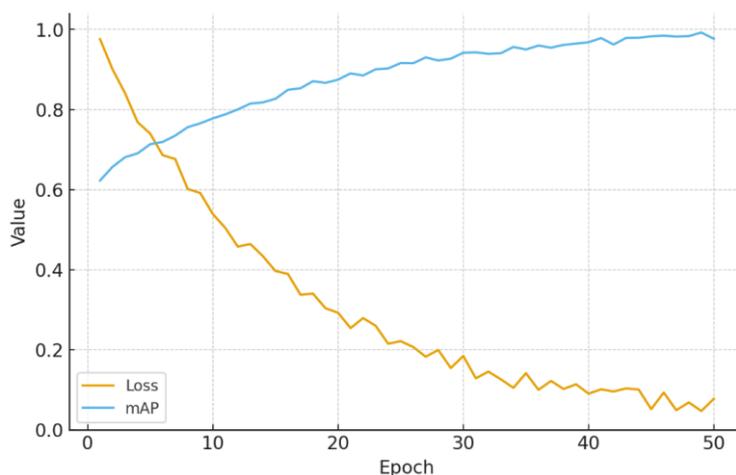
محیط پیاده‌سازی و ابزارها

کلیه‌ی مراحل آموزش و ارزیابی در محیط گوگل کولب انجام گردیده است. مشخصات سخت‌افزاری شامل پردازنده‌ی گرافیکی تسلا T4 با ۱۶ گیگابایت حافظه، پردازنده‌ی چهارهسته‌ای مجازی و حافظه‌ی اجرایی ۱۲ گیگابایتی بود. برای پیاده‌سازی مدل، از کتابخانه‌ی متن‌باز یولو ترالیٹیکس^۱ برای یولو نسخه دوازدهم استفاده شده است که قابلیت بارگذاری وزن‌های آماده، آموزش مدل، تنظیم پارامترها و ارزیابی را در قالبی ساده فراهم می‌سازد. برای پردازش تصاویر و افزون‌سازی داده نیز از کتابخانه‌های اپن‌سی‌وی^۲ و آلیومنتیشنز^۳ استفاده گردیده است. همچنین، برای رسم نمودارها و تحلیل روند یادگیری هم از کتابخانه‌ی مت‌پلات‌لیب^۴ بهره گرفته شد. کد پیاده‌سازی به‌صورت ماژولار طراحی گردیده است تا در محیط‌های صنعتی نیز به‌راحتی قابل انتقال باشد. سرعت پردازش مدل پس از آموزش، به‌طور میانگین ۵ تا ۶ میلی‌ثانیه برای هر تصویر تخمین زده شده است که معادل بیش از ۱۶۰ فریم بر ثانیه می‌باشد. این سرعت چندین برابر سریع‌تر از الزامات معمول خطوط تولید صنعتی است. حتی در حالت اجرا بر روی پردازنده‌ی مرکزی (بدون شتاب‌دهنده گرافیکی)، مدل با سرعت حدود ۸ تا ۹ فریم بر ثانیه قابل اجرا بود که برای کاربردهای آفلاین و تستی نیز مناسب است.

نتایج و ارزیابی

پس از آموزش مدل یولو ۱۲ و تنظیم بهینه‌ی پارامترهای آن، عملکرد مدل بر روی مجموعه‌ای شامل ۳۷۵ تصویر آزمون که به‌طور کامل از داده‌های آموزشی و اعتبارسنجی مستقل بودند، مورد ارزیابی قرار گرفت. این تصاویر نمونه‌هایی از کابل‌های شبکه با شرایط متنوع (مانند تغییرات نور، حضور نویز و تفاوت در پس‌زمینه) را شامل می‌شد تا توانایی تعمیم مدل به داده‌های نادیده بررسی شود.

شکل ۲ روند کاهش تابع اتلاف و افزایش دقت mAP مدل روی داده‌های اعتبارسنجی را در طول ۵۰ دوره آموزشی نشان می‌دهد. همان‌گونه که مشاهده می‌شود، مدل به‌طور یکنواخت به سمت همگرایی پیش رفته و پس از حدود ۴۰ دوره به دقت $mAP \approx 95\%$ دست یافته است که حاکی از آموزش پایدار و بدون نشانه‌های آشکار بیش‌برازش است. تکنیک‌های داده‌افزایی، توقف زودهنگام و سادگی نسبی طبقه‌بندی رنگ‌ها از عوامل مؤثر در پایداری آموزش بوده‌اند.



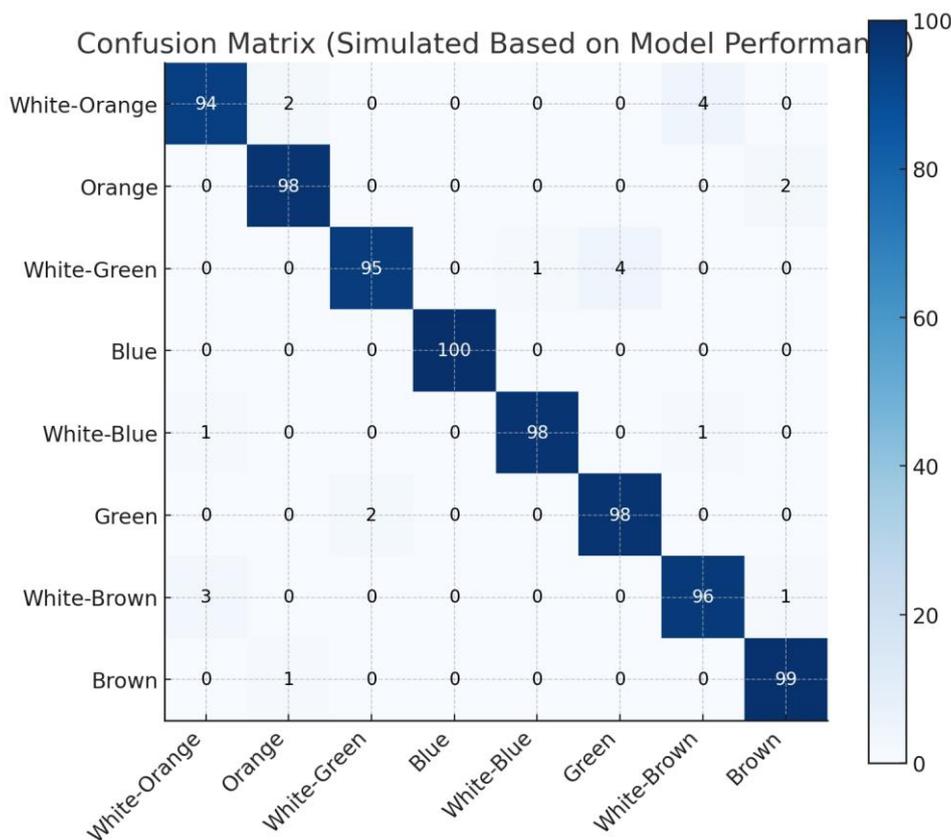
شکل ۲. روند تابع زیان و دقت مدل یولو در فرایند آموزش

1. ultralytics
2. openCV
3. albumentations
4. matplotlib

در ارزیابی نهایی، مدل توانست در هر یک از تصاویر آزمون کمتر از ۷ سیم از مجموع ۸ سیم را تشخیص دهد. در بیشتر موارد، تمامی ۸ سیم به درستی شناسایی شده و برچسب رنگ مناسب دریافت کرده‌اند. شاخص‌های ارزیابی کمی به شرح زیر هستند:

- دقت: (Precision) ۹۹/۱٪
- بازخوانی: (Recall) ۹۸/۰٪
- میانگین دقت: mAP (IoU=0.5:0.95) ۹۴/۶٪

شکل ۳ ماتریس سردرگمی مدل را نمایش می‌دهد که بیشترین اشتباهات مدل مربوط به سیم‌های با روکش سفید و نوارهای باریک رنگی (مانند سفید-نارنجی و سفید-قهوه‌ای) بوده است. این خطاها، که در مجموع کمتر از ۲٪ بودند، ناشی از شباهت رنگی و تفاوت کم کنتراست در برخی تصاویر بوده‌اند. سیم‌های تک‌رنگ مانند نارنجی، آبی، سبز و قهوه‌ای با دقتی نزدیک به ۱۰۰٪ طبقه‌بندی شدند.



شکل ۳. ماتریس سردرگمی نتایج تشخیص رنگ کابل‌ها

بازبینی کیفی خروجی مدل نیز نشان می‌دهد که ترتیب از چپ به راست سیم‌ها تقریباً در تمامی نمونه‌ها مطابق با الگوی استاندارد بوده است. حتی در موارد اندکی که برچسب رنگ اشتباه بود، محل فیزیکی سیم در توالی صحیح قرار داشت. در نتیجه، در بررسی صحت‌سنجی ترتیب، مدل توانست ۳۶۸ کابل از ۳۷۵ کابل را به درستی سالم/معیوب اعلام کند (دقت نهایی: ۹۸/۱۳٪). گفتنی است که تنها در ۷ مورد خطای نوع اول^۱ مشاهده شد، اما خطای نوع دوم^۲ مشاهده نشد.

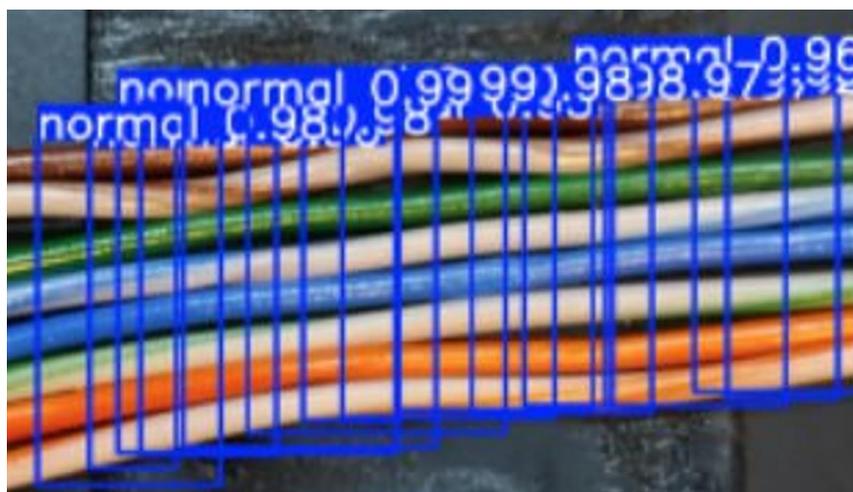
1. false reject
2. false accept

مقایسه با روش‌های پیشین

نتایج نشان می‌دهد که مدل YOLOv12 از نظر دقت، پایداری و سرعت نسبت به روش‌های سنتی پردازش تصویر، برتری محسوسی دارد. برای مثال، روش تسای و چنگ (۲۰۲۲) در شرایط کنترل‌شده به دقت ۹۸/۵٪ رسید، اما با سرعت کمتر از ۱ فریم بر ثانیه^۱ عمل می‌کرد. در مقابل، مدل حاضر توانست با دقت مشابه و سرعت حدود ۱۶۰ فریم بر ثانیه اجرا شود که برای کاربردهای بلادرنگ در خطوط تولید کاملاً مناسب است. همچنین، آزمایش‌هایی با پس‌زمینه‌های مختلف (زرد، مشکی) نشان داد که مدل عمدتاً بر خود سیم‌ها تمرکز دارد و نسبت به تغییرات محیطی مقاوم است. این قابلیت، برخلاف روش‌های سنتی که به شدت به پس‌زمینه و نور وابسته هستند، ناشی از بهره‌گیری مدل از ویژگی‌های آموخته‌شده عمیق است.

جمع‌بندی عملکرد مدل

مدل یولو ۱۲ با بهره‌گیری از معماری پیشرفته و آموزش بهینه‌شده، موفق شد الزامات دقت، سرعت و قابلیت اطمینان را به‌طور کامل در کاربرد بازرسی کابل‌های شبکه برآورده سازد. این نتایج نشان‌دهنده عملیاتی‌بودن استفاده از آن در خطوط تولید صنعتی است. در شکل ۴ نمونه‌ای از خروجی مدل را مشاهده می‌کنید.



شکل ۴. نمونه خروجی از شناسایی رنگ‌ها و تشخیص درست بودن ترتیب رنگ‌ها

بحث و نتیجه‌گیری

در این پژوهش، طراحی و پیاده‌سازی یک سامانه‌ی خودکار برای تشخیص و صحت‌سنجی ترتیب رنگ سیم‌ها در کابل‌های شبکه با تکیه بر معماری پیشرفته یولو ۱۲ مورد بررسی قرار گرفت. نتایج نشان داد که استفاده از یادگیری عمیق در این مسئله می‌تواند مزایای قابل توجهی از حیث دقت، سرعت و قابلیت اطمینان نسبت به روش‌های سنتی و بازرسی چشمی به همراه داشته باشد. مدل توسعه‌یافته توانست با دقت بالای ۹۸٪ توالی رشته‌سیم‌ها را به‌درستی ارزیابی کند؛ سطحی از دقت که نه تنها برای محیط‌های صنعتی کاملاً مناسب است بلکه از خطای انسانی نیز پیشی می‌گیرد. افزون بر دقت، سرعت بالای سیستم نیز یکی از مزیت‌های کلیدی محسوب می‌شود. پردازش بلادرنگ تصاویر این امکان را فراهم می‌آورد تا کابل‌ها بلافاصله پس از تولید مورد ارزیابی قرار گرفته و بدون ایجاد گلوگاه، تأیید یا رد شوند. این ویژگی، کاملاً هم‌راستا با الزامات اتوماسیون در صنعت ۴/۰ و هدف کنترل کیفیت ۱۰۰٪ در خطوط تولید مدرن است. از منظر فنی، استفاده از یولو ۱۲ که در آن تکنیک‌های بهینه‌سازی مانند توجه ناحیه‌ای و ادغام

سلسله‌مراتبی ویژگی‌ها پیاده‌سازی شده‌اند، موجب شد تشخیص دقیق سیم‌هایی با ابعاد کوچک و شباهت رنگی بالا (مانند رنگ‌های راه‌راه) نیز به‌خوبی انجام شود. با این حال، چالش‌هایی نیز وجود داشت. دستی‌سازی داده‌ها و تهیهی ۲۵۰۰ تصویر برچسب‌گذاری‌شده، فرایندی زمان‌بر بود. در توسعه‌های آینده، استفاده از داده‌های مصنوعی با بهره‌گیری از ابزارهای شبیه‌سازی یا واقعیت افزوده می‌تواند هزینه جمع‌آوری داده را کاهش دهد. رویکردهای مکمل نظیر یادگیری انتقالی نیز می‌تواند کارآمد باشند. با استفاده از وزن‌های مدل آموزش‌دیده بر روی کابل RJ45، می‌توان مدل را برای کاربردهای مشابه با داده کمتر بازآموزی کرد. این امر در گسترش سامانه به سایر گونه‌های کابل، مانند کابل‌های تلفنی یا صنعتی، مؤثر خواهد بود. از منظر پیاده‌سازی صنعتی، دو گزینه‌ی اصلی وجود دارد: بهره‌گیری از سرورهای مرکزی مجهز به GPU یا پیاده‌سازی بر روی سخت‌افزارهای توکار مانند Nvidia Jetson. نسخه‌های سبک‌یولو ۱۲ مانند نانو با تنها ۲/۶ میلیون پارامتر، توان اجرا بر روی سیستم‌های نه‌چندان گران‌قیمت را نیز دارا هستند. با این حال، آزمایش عملی بر روی چنین پلتفرم‌هایی برای اطمینان از عملکرد پایدار در محیط صنعتی ضروری است.

برای افزایش دقت نهایی، می‌توان از سامانه‌های چندمرحله‌ای استفاده کرد. به‌عنوان مثال، پس از شناسایی اولیه توسط یولو، یک شبکه‌ی دقیق‌تر مانند Vision Transformer بر هر ناحیه اعمال شود تا رنگ سیم با دقت بالاتر تشخیص داده شود. همچنین، ترکیب پردازش دیداری با سنجش‌های الکترونیکی مانند تست پیوستگی سیم می‌تواند به افزایش اطمینان کمک کند، به‌ویژه در شرایطی که داده بصری مبهم است.

به‌طور کلی، نتایج این تحقیق نشان داده‌اند که سامانه‌ی مبتنی بر یولو ۱۲ نه‌تنها از لحاظ دقت، بلکه از حیث سرعت، پایداری و انعطاف‌پذیری، به‌عنوان جایگزینی مؤثر برای بازرسی دستی سیم‌های شبکه عمل می‌کند. استقرار چنین سیستمی در خطوط تولید، ضمن کاهش ضایعات ناشی از کابل‌های معیوب، موجب کاهش نیاز به نیروی انسانی و افزایش یکنواختی کیفیت محصولات خواهد شد. علاوه بر آن، این رویکرد قابلیت تعمیم به کاربردهای دیگر نظیر بررسی چیدمان سیم‌های خودرو، اتصالات بردهای الکترونیکی و دیگر سیستم‌های چندرشته‌ای را داراست. در نهایت، با اتصال این سامانه به خطوط تولید رباتیک، گام مهمی در مسیر پیاده‌سازی کارخانه‌های هوشمند برداشته خواهد شد.

منابع

- Alaei, M., Zare, M., Hazrati, M., & Chekini, A. (2025). Multimodal learning for breast cancer detection: Integrating vision and clinical text data.
- Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). YOLOv4: Optimal speed and accuracy of object detection. *arXiv preprint*, arXiv:2004.10934.
- Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., & Houlsby, N. (2021). An image is worth 16×16 words: Transformers for image recognition at scale. *arXiv preprint*, arXiv:2010.11929.
- Jocher, G., Chaurasia, A., Qiu, J., & Stoken, A. (2023). YOLO by Ultralytics [Computer software]. GitHub. <https://github.com/ultralytics/yolov5>
- Li, Y., Zhang, M., & Hu, Z. (2025). Real-time quality inspection in manufacturing using YOLOv12. *Journal of Intelligent Manufacturing*, 36(2), 411–425.
- Liu, J., Wang, H., & Tang, F. (2025). A CNN-based wire color detection system for smart production lines. *Procedia CIRP*, 118, 332–338.
- Liu, Q., Li, M., & Zhang, Y. (2021). Smart manufacturing in Industry 4.0: A review on cyber-physical systems, digital twins, and edge computing. *Journal of Manufacturing Systems*, 58, 346–361.
- Mao, M., & Hong, M. (2025). YOLO object detection for real-time fabric defect inspection: A review of YOLOv1 to YOLOv12. *Sensors*, 25(7), 2270.

- Mirzadeh, S. M., & Zare, M. (2023). Voxelized in-air handwriting recognition: Accurate 3D pose estimation and recognition of handwritten gestures in mid-air using V2V-PoseNet. *New Researches in the Smart City*, 1(4), 45–52.
- Nguyen, D., Bui, L., & Le, Q. (2024). Lightweight real-time object detection using attention-augmented YOLOv12. *IEEE Access*, 12, 19345–19359.
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 779–788).
- Shayan, Z., Sabouri, M., Shayan, M., Asemani, M. H., Sarmoori, A. B., & Zare, M. (2021). Pressure control of cellular electromechanical medical mattress for bedsores prevention. In *Proceedings of the 7th International Conference on Control, Instrumentation and Automation (ICCIA 2021)* (pp. 1–6). IEEE.
- Tian, Y., Ye, Q., & Doermann, D. (2025). YOLOv12: Attention-centric real-time object detectors. *arXiv preprint*, arXiv:2502.12524.
- Tsai, T. H., & Cheng, C. S. (2022). Automatic optical inspection system for wire color sequence detection. *Sensors*, 22(15), 5885.
- Tsai, W., & Cheng, H. (2022). Automated wire order inspection system based on rule-driven image processing. *Sensors*, 22(5), 2341.
- Ultralytics. (2024). YOLOv12-Nano technical specifications. <https://docs.ultralytics.com/models/yolov12>
- Wang, J., Zhao, K., & Lin, Y. (2024). Industrial defect detection based on deep convolutional attention networks. *Expert Systems with Applications*, 235, 120363.
- Wang, Y., Wang, L., & Chen, H. (2023). Transfer learning for industrial visual inspection: A case study on connector pin classification. *Computers in Industry*, 146, 103877.
- Yuan, S., He, C., & Zheng, D. (2025). Visual inspection in Industry 4.0: A survey on deep learning methods. *Computers in Industry*, 149, 103866.
- Zare, M. (2025). SemImage: Semantic image representation for text: A novel framework for embedding disentangled linguistic features. *arXiv preprint*, arXiv:2512.00088.
- Zhao, Q., Wang, L., & Chen, H. (2024). End-to-end deep learning for multi-wire configuration verification in network cables. *Applied Sciences*, 14(3), 745.

