تشخیص کلاه ایمنی

گزارش فنی پروژه بر اساس متدولوژی CRISP-DM

نام و نام خانوادگی: محمدحسین دادخواه

تاریخ :مرداد ۱۴۰۵

لینک ریپازتوری گیت‌هاب:

https://github.com/MHDadkhah172/HelmetDitection

1 مقدمه 3
1.1 اهمیت ایمنی در محیط‌های صنعتی 3
1.2 چالش‌های نظارت سنتی..... 3
1.3 هدف پروژه 3
2 ورک‌فلو و ساختار پروژه..... 4
2.1 متدولوژیCRISP-DM 4
3 شناخت داده‌ها (DATA UNDERSTANDING): 5
3.1 مشخصات دیتاست..... 5
4 مشخصات و تغییرات پیش‌فرض دیتاست (DATASET AUGMENTATION FEATURES):..... 6
5 مدل‌سازی و پیکربندی آموزش(MODELING) 7
5.1 علت انتخاب YOLO11 Nano: 7
5.2 انتقال یادگیری (Transfer Learning):..... 7
5.3 جدول هایپرپارامترها (Hyperparameters): 8
6 ارزیابی عملکرد آزمایشگاهی (MODEL EVALUATION): 8
6.1 چالش ناهمگونی داده‌ها (Data Imbalance Analysis):..... 9
6.2 تحلیل شاخص‌های کلیدی 10
7 تست دنیای واقعی و کشف محدودیت‌ها (REAL-WORLD INFERENCE & FAILURE ANALYSIS)..... 10
7.1 کشف خطای وابستگی رنگی(Color Bias) 10
8 پیشنهادات توسعه: معماری دو مرحله‌ای (TWO-STAGE CASCADE ARCHITECTURE)..... 11
9 جمع‌بندی 13

1 مقدمه

1.1 اهمیت ایمنی در محیط‌های صنعتی

تجهیزات حفاظت فردی (Personal Protective Equipment - PPE) و به‌ویژه کلاه ایمنی، اصلی‌ترین ابزار جهت جلوگیری از آسیب‌های فیزیکی شدید و حوادث منجر به فوت در محیط‌های ساختمانی، کارگاهی، پالایشگاهی و معدنی محسوب می‌شوند. بر اساس استانداردهای بهداشت، ایمنی و محیط زیست(HSE) ، حضور تمامی نیروهای کار بدون کلاه ایمنی در محدوده عملیاتی ممنوع است.

1.2 چالش‌های نظارت سنتی

نظارت دستی و چشمی روی رعایت اصول ایمنی توسط ناظران انسانی همیشه کارآمد نیست. این روش با چالش‌های ساختاری زیر روبرو است:

- **خطای انسانی و خستگی:** عدم امکان تمرکز مداوم ناظران در شیفت‌های کاری طولانی.
- **کوری زاویه‌ای (Blind Spots):** عدم پوشش کامل تمام نقاط کارگاه‌های وسیع یا چندطبقه.
- **هزینه‌های بالای پایش:** نیاز به جذب نیروهای نظارتی متعدد و صرف زمان زیاد.

1.3 هدف پروژه

هدف از این تسک، طراحی و توسعه یک سیستم هوشمند بینایی ماشین (Computer Vision) با استفاده از **YOLO11** است تا فرآیند پایش تصاویر ویدئویی را به صورت خودکار و لحظه‌ای (Real-time) انجام دهد. سیستم باید حضور نیروهای کار را بررسی کرده و دو کلاس اصلی را از هم تفکیک کند:

1. helmet: فرد دارای کلاه ایمنی استاندارد (وضعیت ایمن).
2. head: سر بدون کلاه ایمنی (وضعیت خطر / عدم رعایت HSE).

2 ورک‌فلو و ساختار پروژه

2.1 متدولوژیCRISP-DM

برای اجرای مهندسی و ساختاریافته این پروژه، از استاندارد صنعت داده یعنی **CRISP-DM** استفاده شده است. ورک‌فلو پروژه شامل مراحل زیر است:

1. **درک مسئله کسب‌وکار (HSE Needs):** شناسایی چالش‌های پایش دستی ایمنی و تعریف اهداف عملیاتی پروژه.
2. **جمع‌آوری و درک داده‌ها (Roboflow Dataset):** استخراج دیتاست، شناسایی کلاس‌ها و بررسی توزیع نمونه‌ها.
3. **آماده‌سازی داده‌ها (Structure & Augmentation):** اعمال افزونگی داده‌ها (تغییر نور، چرخش) و قالب‌بندی استاندارد فرمت YOLO.
4. **آموزش و بهینه‌سازی مدل (YOLO11n Fine-tuning):** پیکربندی هایپرپارامترها و آموزش شبکه روی پردازنده گرافیکی (GPU).
5. **ارزیابی عمیق آزمایشگاهی (Confusion Matrix & Metrics):** استخراج دقت، فراخوانی، شاخص mAP و تحلیل ماتریس ها حاصل روی داده‌های Test.
6. **استنتاج روی تصاویر دنیای واقعی (Real-World Inference):** تست عملیاتی اسکریپت روی عکس‌های جدید و ناآشنا از محیط‌های کارگاهی.
7. **تحلیل محدودیت‌ها و پیشنهاد معماری آینده:** شناسایی سوگیری رنگی (Color Bias) و ارائه راهکار معماری دو مرحله‌ای (Two-Stage Pipeline).

3 شناخت داده‌ها (Data Understanding):

3.1 مشخصات دیتاست[1]

داده‌های مورد استفاده از دیتاست متن‌باز **Hard Hat Workers** در Roboflow استخراج شده‌اند. داده اولیه شامل ۳ کلاس متغیر بوده است:

1. head: سر بدون کلاه
2. helmet: کلاه ایمنی
3. person: تشخیص کلی پیکر انسان

به عنوان نمونه، در مجموعه داده‌های آزمون (Test Set) از مجموع کل لیبل‌ها، تعداد **۱۹۱۵** نمونه مربوط به کلاس helmet، تعداد **۷۲۶** نمونه مربوط به کلاس head و تنها **۶۴** نمونه مربوط به کلاس person است. این نسبت نشان‌دهندهٔ تمرکز اصلی دیتاست روی تجهیزات ایمنی و سهم بسیار پایین کلاس انسان می‌باشد.

بخش داده‌ها(split)	تعداد تصاویر (images)	درصد از کل
Train Set	۱۴,۷۴۸	۸۸٪
Valid Set	۱,۴۱۳	۸٪
Test Set	۷۰۶	۴٪
مجموع کل	**۱۶,۸۶۷**	**۱۰۰٪**

جدول ۱-۳: مشخصات تفکیک مجموعه‌داده (Dataset Split) و تعداد تصاویر هر بخش

1 https://universe.roboflow.com/joseph-nelson/hard-hat-workers/dataset/11

4 مشخصات و تغییرات پیش‌فرض دیتاست (Dataset Augmentation Features):

مجموعه دادهٔ دریافت‌شده از **Roboflow** جهت افزایش تنوع تصویری، کاهش بیش‌برازش (Overfitting) و بالا بردن توانایی تعمیم‌دهی مدل در شرایط سخت کارگاهی، به صورت پیش‌فرض شامل تکنیک‌های **Data Augmentation** بوده است.

در این دیتاست، به ازای هر تصویر اصلی، چند تصویر جدید با تغییرات تصادفی تولید شده است که مشخصات آن به شرح زیر است:

- **ویژگی‌های هندسی و مکانی (Geometric Features):**
  - **چرخش افقی (Horizontal Flip):** برای شبیه‌سازی جهت‌های مختلف قرارگیری افراد و کلاه‌ها.
  - **برش و زوم (Crop):** زوم تصادفی بین 0% تا 20% جهت شبیه‌سازی فواصل مختلف سوژه تا دوربین.
  - **چرخش زاویه‌ای (Rotation):** چرخش تصادفی تصویر بین -10° تا $+10^\circ$
- **ویژگی‌های رنگی و نوری (Photometric / Color Features):**
  - **سیاه و سفید (Grayscale):** اعمال شده روی 10% تصاویر جهت کاهش وابستگی مدل به رنگ کلاه‌ها (کاهش Color Bias).
  - **تغییر رنگ‌مایه (Hue):** تغییر زاویه رنگ بین -20° تا $+20^\circ$.
  - **غنای رنگ (Saturation):** تغییر غلظت بین −25% تا +25%.
  - **روشنایی و نوردهی (Exposure & Brightness):** تغییرات نور بین −20% تا +20% جهت شبیه‌سازی شرایط نوری متفاوت (سایه، روز و شب).
- **تغییرات کیفی و انسداد (Occlusion Features & Quality):**

- **تاری تصویر (Blur):** اعمال تاری تصادفی تا سقف 1px
- **پوشیدگی و انسداد (Cutout):** وجود ۶ کادر سیاه‌رنگ تصادفی (هر کدام با ابعاد 3% تصویر) روی عکس‌ها برای شبیه‌سازی **انسداد (Occlusion)** در صورت قرار گرفتن سر یا کلاه پشت داربست و موانع کارگاهی.

4-1 دو نمونه از تصاویر دیتاست

5 مدل‌سازی و پیکربندی آموزش(Modeling)

5.1 علت انتخاب YOLO11 Nano:

در سیستم‌های پایش ایمنی صنعتی، **سرعت پردازش بالا (High FPS)** و امکان اجرا روی دستگاه‌های Edge(مانند پردازنده‌های داخلی دوربین‌های مداربسته) اهمیت بالایی دارد. معماری **YOLO11n** به دلیل تعادل فوق‌العاده میان دقت، نرخ فریم بالا و حجم کم مدل انتخاب شد.

5.2 انتقال یادگیری (Transfer Learning):

به جای آموزش مدل از صفر، از وزن‌های پیش‌آموزش‌دیده (Pre-trained) شبکه روی مجموعه داده COCOبهره گرفته شد. این کار باعث شد مدل ویژگی‌های پایه بینایی (لبه‌ها، رنگ‌ها و بافت‌ها) را از قبل شناخته باشد و فرآیند آموزش مستقیماً روی ویژگی‌های تخصصی کلاه و سر متمرکز شود.

5.3 جدول هایپرپارامترها (Hyperparameters):

پارامتر	**مقدار تنظیم‌شده**	**توجیه فنی**
معماری شبکه	YOLO11 Nano	بهینه‌شده برای نرخ فریم بالا و اجرا روی Edge
ابعاد تصویر ورودی	640 × 640	ابعاد استاندارد برای حفظ جزئیات اشیاء کوچک
تعدادEpochs	۱۵	همگرایی مناسب Loss بدون بروز Overfitting
Batch Size	۱۶	بهینه‌سازی بهره‌وری حافظه GPU
محیط پردازشی	Google Colab (GPU T4)	شتاب‌دهی سخت‌افزاری پردازش

6 ارزیابی عملکرد آزمایشگاهی (Model Evaluation):

6.1 تعیین و تنظیم آستانه اطمینان (Confidence Threshold Analysis):

پیش از بررسی شاخص‌های عملکرد و استخراج جداول ارزیابی، آستانه اطمینان (conf) مدل بر اساس **منحنی F1-Confidence**(شکل 1-6) تنظیم گردید. آستانهٔ اطمینان تعیین می‌کند که مدل تنها کادرهایی (Bounding Boxes) را به عنوان خروجی نهایی بپذیرد که درصد اطمینان پیش‌بینی آن‌ها از یک حد مشخص بالاتر باشد.

با تحلیل نقطهٔ اوج منحنی F1-Curve (توازن بهینه میان Precision و Recall)، مقدار آستانه روی conf = 0.35 جهت استنتاج نهایی و محاسبهٔ شاخص‌های ارزیابی تنظیم گردید.

6.2 چالش ناهمگونی داده‌ها (Data Imbalance Analysis):

در فاز ارزیابی روی مجموعه داده Test، مشخص شد کلاس person تنها دارای ۶۴ نمونه است. این حجم کم داده باعث شد مدل نتواند الگوهای بدن انسان را یاد بگیرد ($mAP = 0.0$) و معدل کل تمام کلاس‌ها افت پیدا کند.

از آنجا که هدف اصلی سیستم، **بررسی رعایت ایمنی (کلاه و سر)** است، ارزیابی در دو سطح گزارش می‌شود:

1. نتایج ارزیابی خام (تمامی ۳ کلاس):

کلاس	تعداد نمونه‌ها	Precision	Recall	mAP@50	mAP@50-95
head	۷۲۶	۰.۹۴۸	۰.۹۲۲	۰.۹۳۵	۰.۶۴۹
helmet	۱۹۱۵	۰.۹۵۶	۰.۹۴۰	۰.۹۵۱	۰.۶۶۰
person	۶۴	۰.۰۰۰	۰.۰۰۰	۰.۰۰۰	۰.۰۰۰
معدل کل (All Classes)	**۲۷۰۵**	**۰.۶۳۵**	**۰.۶۲۱**	**۰.۶۲۹**	**۰.۴۳۶**

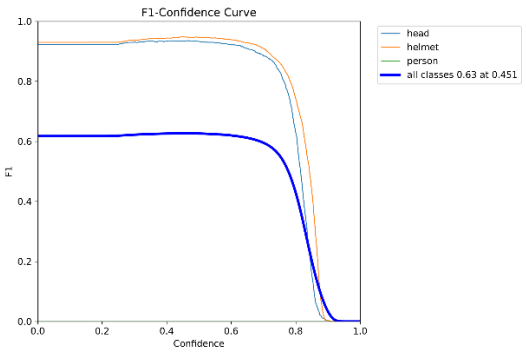
2. نتایج ارزیابی واقعی (کلاس‌های اهداف ایمنی):

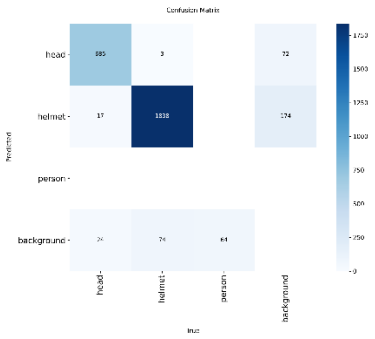
با حذف کلاس کم‌تعداد person و تمرکز بر اهداف اصلی (helmet , head) عملکرد فوق‌العاده مدل شفاف می‌شود:

کلاس	تعداد نمونه‌ها	Precision	Recall	mAP@50	mAP@50-95
head	۷۲۶	۰.۹۴۸	۰.۹۲۲	۰.۹۳۵	۰.۶۴۹
helmet	۱۹۱۵	۰.۹۵۶	۰.۹۴۰	۰.۹۵۱	۰.۶۶۰
معدل اهداف	**۲۶۴۱**	**۰.۹۵۲**	**۰.۹۳۱**	**۰.۹۴۳**	**۰.۶۵۵**

6.3 تحلیل شاخص‌های کلیدی

- **دقت بالا روی اهداف اصلی ($mAP@50 = 94.3\%$):** مدل با ثبت این امتیاز، بسیار بالاتر از حد نصاب موفقیت اولیه (۸۰٪) قرار گرفت.
- **مرزبندی دقیق Bounding Boxها ($mAP@50-95 = 65.5\%$):** نشان‌دهنده انطباق بسیار دقیق کادرهای پیش‌بینی‌شده با مرزهای سر و کلاه است.
- **تحلیل ماتریس اغتشاش (Confusion Matrix):** مدل از مجموع ۷۲۶ نمونه سر بدون کلاه، تنها در ۳ مورد دچار خطای اشتباه با کلاه شده است. این نرخ بسیار پایینِ خطا برای سیستم‌های ایمنی حیاتی است.





شکل 6-1 F1 Curve (سمت راست) | شکل 6-2 confusion_matrix (سمت چپ)

7 تست دنیای واقعی و کشف محدودیت‌ها (Real-World Inference & Failure Analysis)

در فاز ۶ تسک، اسکریپت src/predict.py روی تصاویری واقعی و دیده‌نشده اجرا شد.

7.1 کشف خطای وابستگی رنگی(Color Bias)

تحلیل کیفی خروجی‌ها روی تصاویر واقعی یک رفتار محدودکننده را آشکار ساخت:

- **مشکل مشاهداتی:** مدل روی کلاه‌های ایمنی **سفید رنگ**(شکل 7-1) نسبت به کلاه‌های زرد و آبی، آستانه اطمینان (Confidence) کمتری نشان می‌دهد یا در مواردی آن‌ها را نادیده می‌گیرد.

- **علت ریشه‌ای (Spurious Correlation):** به دلیل فراوانی بیشتر کلاه‌های زرد و آبی در داده‌های آموزشی، شبکه بخشی از مفهوم «کلاه» را بارنگ گره زده است.
- **راهکار اصلاحی:** در آموزش‌های بعدی باید از تکنیک **Grayscale Augmentation** استفاده شود تا بی‌رنگ کردن داده‌ها شبکه را مجبور کند صرفاً روی اشکال هندسی، لبه‌ها و بافت کلاه متمرکز شود.

1-7 نمونه‌هایی از تصاویر پردازش‌شده توسط مدل؛ در هر دو تصویر، مدل قادر به تشخیص صحیح کلاه ایمنی سفید رنگ نبوده و دچار خطا شده است.

8 پیشنهادات توسعه: معماری دو مرحله‌ای (Two-Stage Cascade Architecture)

برای استقرار صنعتی این سیستم و به صفر رساندن خطای مهلک (یعنی ندیدن یک کارگر بدون کلاه)، استفاده از به روش زیر به‌نظر بهتر است:

معماری پایپ‌لاین دو مرحله‌ای (Two-Stage Cascade Architecture):

1. **ورودی سیستم:** دریافت فریم‌های تصویری یا جریان ویدئویی زنده از دوربین مداربسته.
2. **مرحله ۱ — مدل تشخیص انسان (Person Detector):**

- **تنظیمات:** آستانه اطمینان پایین
- **هدف اصلی:** اولویت با **High Recall** است تا تحت هر شرایطی (نور کم، فاصله زیاد، اختفای نسبی)، تمام نیروهای کار موجود در تصویر شناسایی شوند و هیچ کارگری از دید سیستم پنهان نماند.
- **خروجی:** برش خودکار ناحیه بدن/سر نیروهای کار (Crop Person Region).

3. **مرحله ۲ — مدل طبقه‌بندی کلاه و سر (Helmet/Head Classifier):**

- **تنظیمات:** آستانه اطمینان بالا
- **هدف اصلی:** اولویت با **High Precision** است؛ یعنی بررسی دقیق نواحی برش‌خورده تا اجسام خارجی پیش‌زمینه (مثل کپسول آتش‌نشانی، دیوار سفید یا تجهیزات) به اشتباه به عنوان کلاه ایمنی شناسایی نشوند.
- **خروجی نهایی:** ثبت وضعیت ایمنی فرد (helmet یا head) و ارسال هشدار در صورت عدم رعایت HSE.

در ارزیابی سیستم‌های ایمنی صنعتی (HSE)، موازنه میان خطاهای مدل همیشه به نفع حفظ جان نیروی کار انجام می‌شود. به بیان دقیق‌تر، وقوع **خطای مثبت کاذب (False Positive)** به مراتب **پذیرفته‌شده‌تر** از **خطای منفی کاذب (False Negative)** است.

در این رویکرد، اگر سیستم به اشتباه یک کارگرِ دارای کلاه را به عنوان فردی بدون کلاه طبقه‌بندی کند (افزایش هشدار کاذب)، خطای مهلکی رخ نداده است؛ اما اگر یک **انسان بدون کلاه به اشتباه دارای کلاه تشخیص داده شود (False Safety Classification)**، یک ریسک بحرانی و صدمه جانیِ جبران‌ناپذیر به همراه خواهد داشت.

9 جمع‌بندی

در این تسک، پایپ‌لاین هوشمند پایش کلاه ایمنی بر اساس متدولوژی **CRISP-DM** و با تکیه بر معماری **YOLO11** پیاده‌سازی و ارزیابی شد. نتایج کلیدی پروژه را می‌توان در سه محور زیر خلاصه کرد:

- **عملکرد بالا در اهداف ایمنی:** مدل نهایی پس از حذف نویز داده‌های غیرمرتبط، به شاخص **mAP@50 = 94.3%** و **Recall = 93.1%** روی کلاس‌های helmet و head دست یافت که نشان‌دهنده توانایی بالای آن در تفکیک سر بدون کلاه از کلاه ایمنی است.
- **شناسایی رفتارهای عملیاتی:** ارزیابی مدل روی داده‌های واقعی، وجود وابستگی رنگی (Color Bias) روی کلاه‌های سفید را آشکار ساخت که ریشه در توزیع داده‌های ورودی داشت.
- **نقشه راه توسعه صنعتی:** برای رفع محدودیت‌های موجود و به صفر رساندن خطای مهلک (Missed Head Detection)، دو راهکار عملیاتی **Grayscale Augmentation** و **پایپ‌لاین دو مرحله‌ای (Two-Stage Cascade)** برای ارتقا به محصول صنعتی پیشنهاد شد.