شناسایی افراد مبتنی بر یادگیری عمیق با منابع محدود با استفاده از ربات چشم چابک

دانشجو: عرفان باقری سولا

استاد راهنما: دکتر طالع ماسوله

دانشکده مهندسی برق و کامپیوتر، دانشگاه تهران

مقدمه

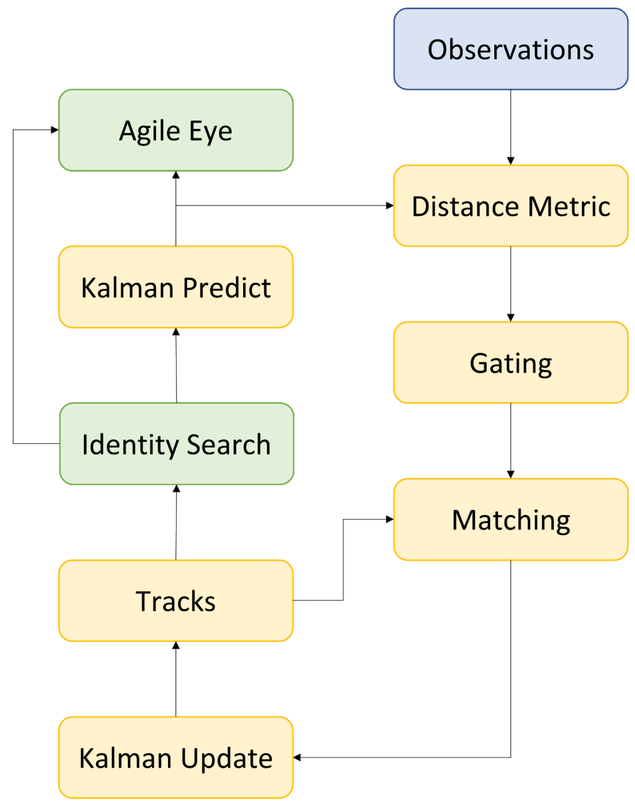
در سال‌های اخیر، گسترش سیستم‌های نظارتی و کاربردهای تعامل انسان و ربات، علاقه رو به رشدی به الگوریتم‌های ردیابی و بازشناسی هویت افراد به وجود آورده است. همچنین روندی جهت انتقال محاسبات از سرورهای متمرکز به دستگاه‌های لبه در حوزه تکنولوژی شکل گرفته است.

در این پژوهش:

- یک سیستم ردیابی اشیاء و بازشناسی هویت افراد با قابلیت اجرای بلادرنگ در سیستم‌هایی که فاقد پردازنده‌های گرافیکی هستند، توسعه داده شده است.
- ربات چشم چابک جهت ردیابی هویت‌های هدف در یک میدان دید وسیع‌تر به کار گرفته شده است.
- عملکرد مدل‌های کوچک‌تر در الگوریتم ردیابی با تنظیم دقیق بهبود داده شده است.

ساختار ردیاب ارائه شده

الگوریتم ردیابی و بازشناسی هویت افراد در این پروژه بر مبنای الگوریتم ردیابی اشیاء Deep-SORT (بلوک‌های زرد رنگ) توسعه داده شده است (شکل ۱).



شکل ۱ – الگوریتم کلی سیستم.

مشاهدات ورودی الگوریتم از دو بخش تشکیل می‌شود:

- موقعیت افراد تشخیص داده شده (به فرم مختصات جعبه مرزی).
- توصیفگر ظاهری افراد تشخیص داده شده (به فرم بردار ویژگی).

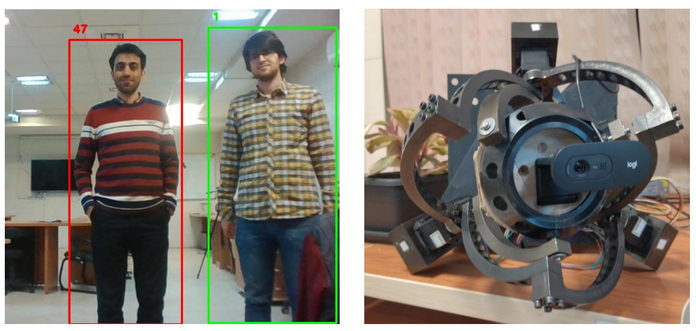
نوآوری‌های این پژوهش به شرح زیر می‌باشد:

- تنظیم دقیق مدل‌های سبک‌وزن تشخیص اشیاء YOLOv8n و YOLOv8s روی مجموعه داده CrowdHuman جهت بهبود عملکرد.
- استفاده از شبکه‌های بهینه OSNet جهت بازشناسی هویت که ۱۰ برابر سبک تر از مدل‌های مبتنی بر ResNet هستند.
- استفاده از شاخص جدایی‌پذیری متقابل جهت ارزیابی مدل‌های بازشناسی هویت.
- اضافه کردن قابلیت تعریف هویت‌های از پیش تعیین شده به الگوریتم Deep-SORT.
- دنبال کردن هویت هدف در یک میدان دید وسیع‌تر با استفاده از ربات چشم چابک.
- استفاده از موقعیت پیش‌بینی شده هویت هدف توسط فیلتر کالمن جهت حرکت دادن ربات چشم چابک برای حداکثر سازی احتمال مشاهده هویت هدف در فریم آینده.
- ثبت هویت‌های شناسایی شده و ارائه گزارش پس از اتمام الگوریتم.

ارزیابی الگوریتم ردیابی

جدول ۱ - عملکرد الگوریتم ردیابی با مدل‌های مختلف روی مجموعه داده MOT16.

Detector	ReID	MOTA ↑	MOTP ↑	IDs ↓
YOLOv8n	OSNet x0.25	42.7 %	79.8 %	626
YOLOv8n FT	OSNet x0.25	55.9 %	79.9 %	605
YOLOv8n FT	OSNet x1.0	55.8 %	79.9 %	**498**
YOLOv8s FT	OSNet x1.0	57.0 %	**80.2 %**	509
Original Deep-SORT Paper		**61.4 %**	79.1 %	781



شکل ۲ – نمونه‌ای از عملکرد الگوریتم ردیابی و ربات چشم چابک.

نتایج به دست آمده از ارزیابی الگوریتم به شرح زیر است:

- ٪۳۱ بهبود در عملکرد الگوریتم ردیابی نسبت به مدل پایه، با تنظیم دقیق مدل تشخیص اشیاء برای ۴۰ ایپاک و به مدت ۱۲ ساعت (جدول ۱).
- کاهش مجموع تعداد پارامترهای مدل‌های استفاده شده به کمتر از ۹ میلیارد پارامتر (در مقایسه با ۱۴۰ میلیارد پارامتر مقاله Deep-SORT) به ازای:
  - ٪۹ کاهش در صحت الگوریتم ردیابی.
  - ثابت ماندن دقت الگوریتم ردیابی.
  - ٪۲۳ بهبود در تعداد شناسه‌های اختصاص داده شده.
- قابلیت اجرا روی پردازنده Intel Core I7 6600u با سرعتی در حدود ۵ فریم بر ثانیه و ردیابی ۵ فرد به صورت همزمان در هر فریم.
- عملکرد مطلوب ربات چشم چابک در سناریوهای تست شده.

جمع بندی

در این پژوهش:

- عملکرد مدل‌های کوچک‌تر در الگوریتم ردیابی بررسی شده‌اند.
- مدل‌های تشخیص اشیاء با هدف بهبود عملکرد در الگوریتم ردیابی افراد، تنظیم دقیق شده‌اند.
- شاخص جدایی‌پذیری متقابل جهت ارزیابی و انتخاب مدل بازشناسی هویت مورد استفاده قرار گرفته است.
- قابلیت تعریف هویت‌های از پیش تعیین شده به Deep-SORT اضافه شده‌اند.
- از ربات چشم چابک جهت دنبال کردن هویت‌های هدف استفاده شده است.

مراجع اصلی

1. Wojke, Nicolai, Bewley, Alex, and Paulus, Dietrich. Simple online and realtime tracking with a deep association metric. In 2017 IEEE International Conference on Image Processing (ICIP).
2. Saffar, Mohsen and Kalhor, Ahmad. Evaluation of dataflow through layers of convolutional neural networks in classification problems. Expert Systems with Applications, 2023.
3. Bonev, Ilian A., Chablat, Damien, and Wenger, Philippe. Working and assembly modes of the agile eye. Proceedings 2006 IEEE International Conference on Robotics and Automation. ICRA 2006.
4. Zhou, Kaiyang, Yang, Yongxin, Cavallaro, Andrea, and Xiang, Tao. Omni-scale feature learning for person re-identification. In 2019 IEEE/CVF International Conference on Computer Vision (ICCV).