گرفتن اجسام به کمک ربات سه انگشتی با استفاده از الگوریتم‌های یادگیری عمیق

دانشجو: فاطمه نائینیان

استاد راهنما: دکتر مهدی طالع ماسوله

دانشکده مهندسی برق و کامپیوتر، دانشگاه تهران

مقدمه

گرفتن اجسام یک جنبه اساسی از تعامل انسان با محیط است و ایجاد این توانایی در سیستم‌های رباتیکی به دلیل پیچیدگی تعاملات دست و شی چالش برانگیز است. انسان‌ها به صورت غریزی به بهترین اجسام را می‌گیرند و این انگیزه‌ای برای آموزش ربات‌ها با رفتار انسانی است تا ربات‌ها بتوانند با کیفیت بیشتر این فعالیت را انجام دهند.

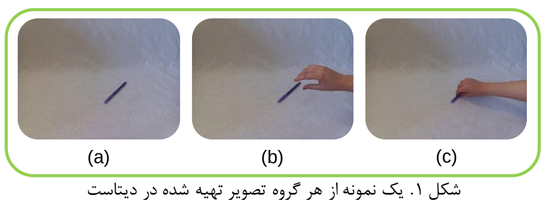
در این پژوهش

- به بررسی نگاشت رفتار انسان در گرفتن اجسام به یک ربات سه انگشتی با استفاده از رویکرد یادگیری عمیق پرداخته می‌شود.
- یک مجموعه داده از رفتار انسان در گرفتن اشیاء تهیه می‌شود.
- مدلی آموزش داده می‌شود که نقاط گرفتن را بر اساس تصاویر رنگی از اشیاء و مختصات انگشت‌ها در حالتی که به جسم نزدیک شده است را پیش‌بینی می‌کند.
- نقطه اشتراک انسان و ربات در این پژوهش، موقعیت انگشتان است که باعث می‌شود تا بتوان از مدلی که با داده‌های انسانی آموزش می‌بینند، برای ربات‌ها نیز استفاده کرد.

روش پیشنهادی

برای ایجاد این نگاشت، مجموعه داده‌ای از نحوه گرفتن اجسام با ویژگی‌های زیر تولید شده است.

- جمع آوری داده از۱۷ شیء در جهت‌ها و سناریوهای مختلف
- ۸۰۱ گروه تصویر با تعداد کل ۲۴۰۳ تصویر
- هر گروه تصویر شامل
  - تصاویر از اشیاء مورد نظر (a)
  - تصاویر نزدیک شدن دست به اشیاء (b)
  - تصاویر در لحظه گرفتن اشیاء (c)



شکل ۱. یک نمونه از هر گروه تصویر تهیه شده در دیتاست

پردازش داده‌ها:

- استخراج موقعیت انگشتان در تصاویر نزدیک شدن دست به اشیاء و تصاویر در لحظه گرفتن اشیاء با مدیاپایپ
- تغییر اندازه تصاویر و برش آنها به اندازه نهایی (۸۰ و ۸۰)
- افزایش داده‌ها شامل چرخش تمام تصاویر و برچسب‌های مربوطه با سه زاویه مختلف و تشکیل مجموعه داده نهایی شامل ۳۲۰۴ گروه سه تایی تصویر

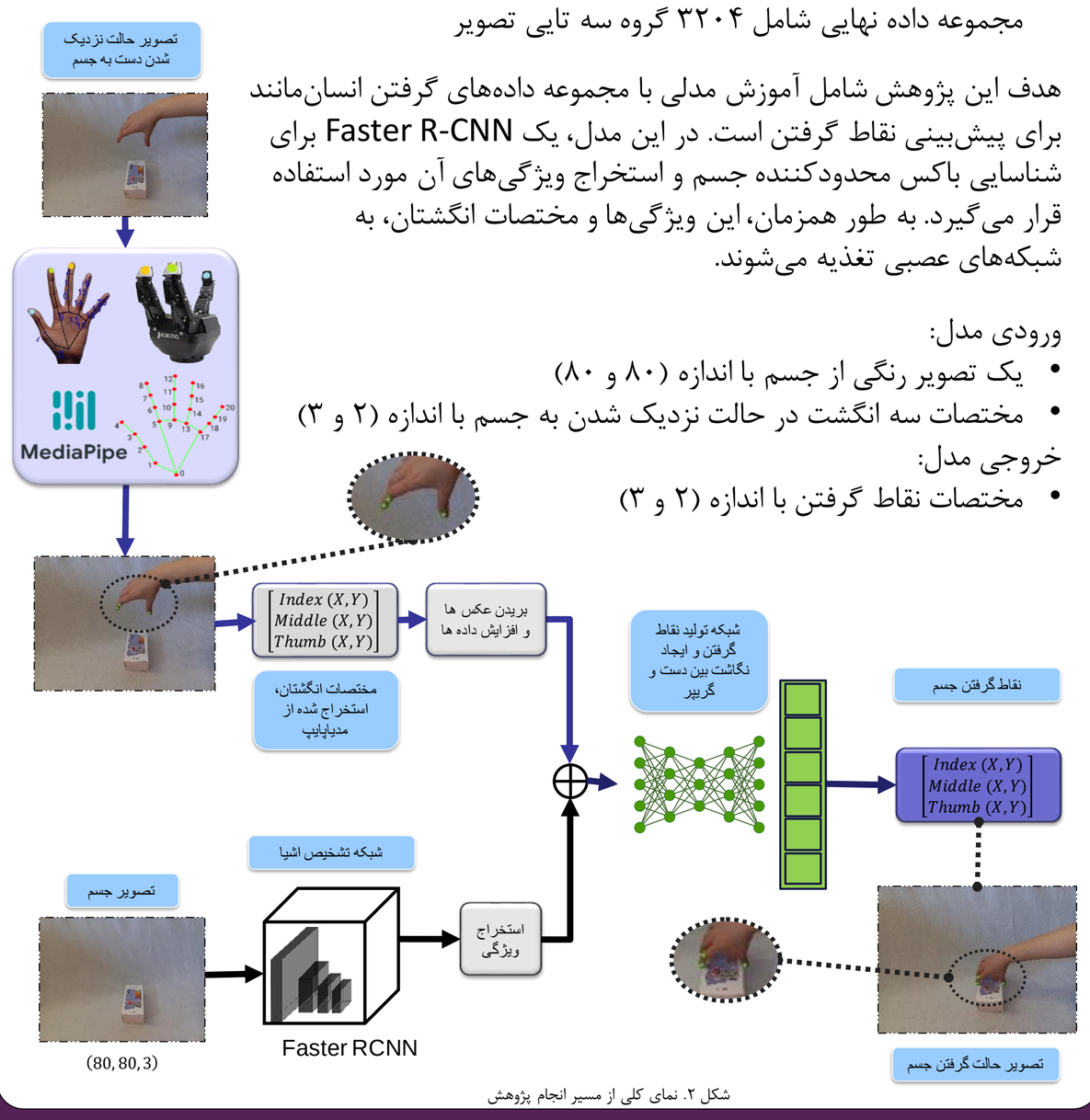
هدف این پژوهش شامل آموزش مدلی با مجموعه داده‌های گرفتن انسان‌مانند برای پیش‌بینی نقاط گرفتن است. در این مدل، یک Faster R-CNN برای شناسایی باکس محدودکننده جسم و استخراج ویژگی‌های آن مورد استفاده قرار می‌گیرد. به طور همزمان، این ویژگی‌ها و مختصات انگشتان، به شبکه‌های عصبی تغذیه می‌شوند.

ورودی مدل:

- یک تصویر رنگی از جسم با اندازه (۸۰ و ۸۰)
- مختصات سه انگشت در حالت نزدیک شدن به جسم با اندازه (۲ و ۳)

خروجی مدل:

- مختصات نقاط گرفتن با اندازه (۲ و ۳)

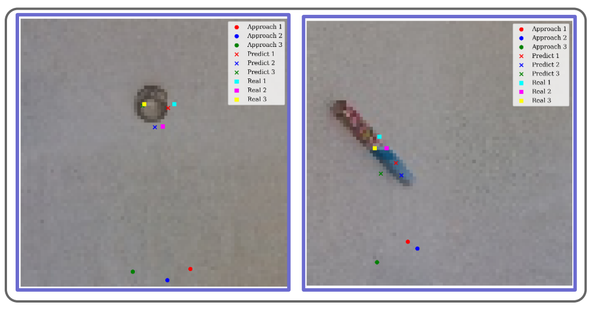


شکل ۲. نمای کلی از مسیر انجام پژوهش

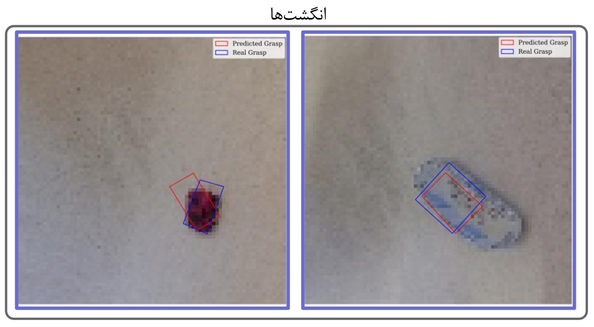
نتایج

این مدل برای ۵۰ ایپاک با استفاده از بهینه ساز Adam آموزش داده شد. تابع ضرر انتخابی نیز میانگین خطای مطلق است. معیارهای ارزیابی مدل در این پژوهش به شرح زیر است:

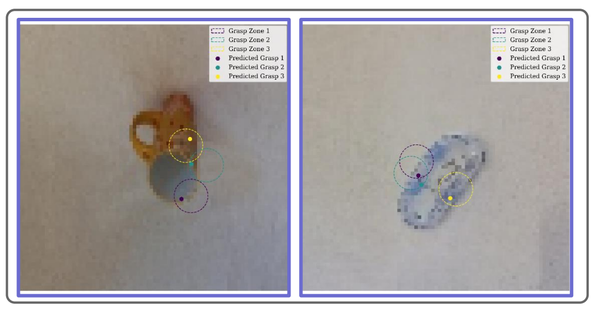
- معیار اشتراک بر اجتماع (IoU): همپوشانی بین منطقه پیش‌بینی‌شده گرفتن جسم و ناحیه حقیقی گرفتن جسم را محاسبه می‌کند.
- زاویه گرفتن: نشان دهنده این است که دست یا گریپر در زاویه مناسبی برای گرفتن یک شی هستند یا خیر. اختلاف زاویه کمتر از ۳۰ درجه نشان دهنده موفقیت پیش‌بینی است.
- دقت پیکربندی گرفتن: میزان امکان پذیری گرفتن یک جسم را با در نظر گرفتن اندازه انگشتان (w) تعیین می‌کند و اگر فاصله بین مختصات پیش‌بینی شده و حقیقی کمتر از $\frac{w}{2}$ باشد، نشان می‌دهد که انگشت می‌تواند با موفقیت جسم را در آن موقعیت بگیرید.



شکل ۳. نقاط گرفتن پیش بینی شده در کنار نقاط گرفتن اصلی و حالت اولیه موقعیت انگشت‌ها



شکل ۴. محاسبه IoU با نمایش مستطیلی



شکل ۵. ناحیه انگشتان برای گرفتن جسم برای معیار دقت پیکربندی گرفتن

جدول 1. نرخ موفقیت تولید نقاط گرفتن جسم

اجسام	معیارها		
	دقت پیکربندی گرفتن	اشتراک بر اجتماع	زاویه گرفتن
جعبه عینک	۸۹.۴۷	۴۹.۱۲	۹۷.۳۶
عروسک جغد	۹۶.۲۶	۳۸.۲۵	۱۰۰
توپ	۸۱.۸۱	۴۴.۹۱	۹۰.۹

جدول 1، نتایج داده‌های آزمون را با معیارهای ارزیابی برای سه شیء نادیده نشان می‌دهد. همانطور که مشاهده می‌شود، مدل قادر است نقطه گرفتن را با دقت بالایی در زاویه گرفتن و دقت پیکربندی گرفتن پیش‌بینی کند. علاوه بر این، توپ، اسباب بازی جغد و جعبه عینک دارای IoU قابل قبولی هستند. در تصاویر ۳، ۴ و ۵ مشاهده می‌شود که مدل پیش‌بینی بسیار خوبی را عرضه می‌کند.

جمع بندی

- در این پژوهش روشی برای پیش‌بینی نقاط گرفتن بر اساس رفتار انسان در گرفتن اشیاء با تصاویر اجسام و مختصات انگشتان در حالت نزدیک شدن به جسم ارائه شده است.
- در این کار تاثیر موقعیت دست یا گریپر هنگام نزدیک شدن به جسم بر روی مختصات نهایی گرفتن، نشان داده شده است.
- این روش نوآورانه آغازی برای آموزش گریپرها با رفتارهای انسانی و بهبود عملکرد آنهاست.
- این روش در صنعت این امکان را می‌دهد تا تعامل انسان و ربات آسان تر شود و ربات‌ها عملکردی مشابه انسان‌ها در گرفتن اجسام انجام دهند.
- این روش با داشتن امکان آموزش با داده‌های انسانی، انسان را از جمع‌آوری داده به صورت مستقیم از ربات یا شبیه‌سازی که بسیار هزینه بر و زمان بر است بی‌نیاز می‌کند.

مراجع اصلی

1. Amirul Sadun, Jamaludin Jalani, and Faizal Jamil. (2016). Grasping analysis for a 3-Finger Adaptive Robot Gripper. 1-6. 10.1109/ROMA.2016.7847806.
2. Wang Dexin. (2020). SGDN: Segmentation-Based Grasp Detection Network For Unsymmetrical Three-Finger Gripper.
3. Elnaz Balazadeh, Mehdi Tale Masouleh, and Ahmad Kalhor. (2023). HUGGA: Human-like Grasp Generation with Gripper's Approach State Using Deep Learning. International Conference on Robotics and Mechatronics.