

دانشگاه صنعتي امیرکبیر  
(پلی تکنیک تهران)

دانشکده مهندسی کامپیوتر

گزارش کارآموزی

موضوع کارآموزی:

**Video Swin Transformer**

|  |  |
| --- | --- |
| نام استاد کارآموزی: | دکتر صاحب الزمانی |
| نام دانشجو: | محمدرضا قادری عمودیزج |
| شماره دانشجویی: | 9627057 |

**تابستان 1399**

**تقدیر و تشکر**

بدینوسیله مراتب قدردانی و امتنان خود را خدمت،

جناب آقای مهندس حسین یعقوب زاده، سرپرست تیم توسعه دهنده پژوهشگاه توسعه فناوری‌های پیشرفته، بابت آموزش‎ها و زحمات بی‌دریغشان، تمامی کارکنان و همراهان بنده در پژوهشگاه توسعه فناوری‌های پیشرفته ،

ابراز و از تمامی زحمات آنان تشکر می‌نمایم.

محمدرضا قادری عمودیزج

شهریور 1400

چكيده

امروزه با توجه به نیازهای جامعه از نظر امنیتی و توسعه هوش مصنوعی و همینطور شناسایی رفتارهای افراد در جامعه از نظر بازدید ویدیوها و تصاویر در شبکه‌های اجتماعی و حتی استفاده از شناسایی این رفتارها در ربات‌های انسان نما برای پاسخ به حرکات پیش آمده نیاز به یادگیری و تشخیص رفتارهای تصاویر و ویدیو شده است. در اواخر قرن بیستم با توجه به رونق هوش مصنوعی از این تکنولوژی برای بررسی و تجزیه وتحلیل خودکار مبنی بر تصویربرداری استفاده شد که بینایی ماشین[[1]](#footnote-1) نامیده شد، که روز به روز در حال گسترش هست.

در این گزارش، پس از تحلیل این روش برای بررسی ویدیوها به صورتی که بتوان با بینایی ماشین به تشخیص نوع رفتاری آنها بیانجامد استفاده می‌کنیم، و نوع رفتاری و انجام تحلیل ویدیوهای، نحوه‌ی استفاده از دیتابیس و ابزارها و کتابخانه‌های لازم رو بررسی می‌کنیم.

واژه‌های کلیدی:

هوش مصنوعی، پردازش تصویر[[2]](#footnote-2)، بینایی ماشین، تشخیص رفتار ویدیو.

فهرست مطالب

[فصل 1: مقدمه](#_Toc69291746) 0

[1-1- مقدمه 0](#_Toc69291747)

[1-2- کارهای مرتبط 0](#_Toc69291748)

[1-3- معماری شبکه 0](#_Toc69291749)

[1-4- جمع‏بندی 0](#_Toc69291750)

[فصل 2: مجموعه دادگان 1](#_Toc69291746)

[2-1- مقدمه 2](#_Toc69291747)

[2-2- نحوه دسترسی 2](#_Toc69291748)

[2-3- جزئیات مجموعه دادگان 2](#_Toc69291749)

[2-4- جمع‏بندی 2](#_Toc69291750)

[فصل 3: راه‏اندازی کد به صورت برخط (Online) 3](#_Toc69291751)

[3-1- مقدمه 4](#_Toc69291752)

[3-2- نصب پکیج‏ها و پیش‏نیازها 4](#_Toc69291753)

[3-3- نحوه دسترسی 4](#_Toc69291754)

[3-4- راه‏اندازی سریع 4](#_Toc69291755)

[3-5- بررسی جزئیات 5](#_Toc69291756)

[3-6- جمع‏بندی 6](#_Toc69291757)

[فصل 4: راه‏اندازی کد به صورت برون خط (Offline) 7](#_Toc69291758)

[4-1- مقدمه 8](#_Toc69291759)

[4-2- نحوه دسترسی 8](#_Toc69291760)

[4-3- نصب پیش‏نیازها به صورت Offline 8](#_Toc69291761)

[4-4- راه‏اندازی سریع 8](#_Toc69291762)

[4-5- بررسی جزئیات کد 8](#_Toc69291763)

[4-6- جمع‏بندی 9](#_Toc69291764)

[فصل 5: جمع‏بندی و پیشنهادات 12](#_Toc69291765)

[5-1- مقدمه 13](#_Toc69291766)

[5-2- جمع‏بندی کلی و پیشنهادات 13](#_Toc69291767)

[5-3- کارهای آینده 13](#_Toc69291768)

[5-4- جمع‏بندی 13](#_Toc69291769)

[مراجع 14](#_Toc69291770)

فهرست اشکال

[شکل (1-1) تصویر نمونه 2](#_Toc34822082)

**فهرست جداول**

[جدول (1-1) جدول نمونه 2](#_Toc34821840)

1. مقدمه
   1. مقدمه
   2. بینایی ماشین و پردازش تصویر
   3. ترانسفورماتور چرخش تصویری[[3]](#footnote-3)
   4. جمع بندی
2. معرفی محل کارآموزی
3. فعالیت‌ها و تجارب کارآموزی
   1. مقدمه
4. جمع بندی و پیشنهادات
   1. مقدمه

متن.

* 1. جمع‏بندی کلی و پیشنهادات

متن.

* 1. کارهای آینده

متن.

* 1. جمع‏بندی

متن.

مراجع

[1] Q. Liu, A. H. Sung, and M. Qiao, "Video steganalysis based on the expanded markov and joint distribution on the transform domains detecting msu stegovideo," in *Machine Learning and Applications, 2008. ICMLA'08. Seventh International Conference on*, 2008: IEEE, pp. 671-674.

1. Machine vision [↑](#footnote-ref-1)
2. Image Processing [↑](#footnote-ref-2)
3. Video Swin Transformer [↑](#footnote-ref-3)