

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



دانشکده مهندسی

گروه مهندسی کامپیوتر

## فیلتر تصاویر غیر اخلاقی بر اساس تشخیص پوست

پایان نامه کارشناسی ارشد

نگارنده

علی ترابیان دهکردی

استاد راهنما

آقای دکتر سید عنایت الله علوی

استاد مشاور

آقای دکتر علیرضا عصاره

شهریور ۱۳۹۳

باسمه تعالی

دانشگاه شهید چمران اهواز  
دانشکده مهندسی

(نتیجه ارزشیابی پایان نامه ارشد)

پایان نامه آقای علی ترابیان دهکردی دانشجوی رشته: مهندسی کامپیوتر گرایش: هوش مصنوعی

دانشکده مهندسی به شماره دانشجویی ۹۱۱۵۴۰۱

با عنوان:

فیلتر تصاویر غیراخلاقی براساس تشخیص پوست

جهت اخذ مدرک: کارشناسی ارشد در تاریخ: ۱۳۹۳/۶/۳۱ توسط هیأت داوران مورد ارزشیابی قرار گرفت و با درجه خیلی خوب تصویب گردید.

امضاء	رتبه علمی	اعضای هیأت داوران:
.....	استادیار	استاد راهنما: دکتر سیدعنایت الله علوی
.....	دانشیار	استاد مشاور: دکتر علیرضا عصاره
.....	استادیار	استاد داور: دکتر احسان نامجو
.....	استادیار	استاد داور: دکتر کریم انصاری اصل
.....	استادیار	نماینده تحصیلات تکمیلی: دکتر علی حقیقی
.....	استادیار	۲. مدیر گروه: دکتر مرجان نادران
.....	استادیار	۳. معاون پژوهشی و تحصیلات تکمیلی دانشکده: دکتر علی حقیقی
.....	استاد	۴. مدیر تحصیلات تکمیلی دانشگاه: دکتر عبدالرحمن راسخ

## فهرست

۱ فصل اول: مقدمه .....	۱
۱-۱ تعریف مسئله .....	۱
۲-۱ انواع روشه‌ای فیلترینگ اینترنت .....	۲
۱-۲-۱ فیلترینگ بر اساس لیست سیاه IP .....	۴
۲-۲-۱ فیلترینگ بر اساس محتوای متنی .....	۴
۳-۲-۱ فیلترینگ بر اساس محتوای دیداری .....	۵
۳-۱ اهداف پایان‌نامه .....	۶
۴-۱ روش استفاده‌شده .....	۷
۵-۱ ساختار پایان‌نامه .....	۸
۲ فصل دوم: پیشینه‌ی تحقیق .....	۹
۳ فصل دوم: مبانی تحقیق .....	۱۶
۱-۳ بسته‌ی کلمات بصری .....	۱۸
۱-۱-۳ تفاوت‌های کلمات بصری و نوشتاری .....	۲۱
۲-۱-۳ مدل بازیابی اطلاعات .....	۲۹
۲-۳ ویژگی‌های محلی .....	۳۳
۱-۲-۳ تشخیص‌دهنده و توصیفگر محلی ویژگی مقاوم شتاب گرفته .....	۳۴
۲-۲-۳ تشخیص‌دهنده مناطق حدی حداکثر پایدار .....	۳۷
۳-۳ تحلیل رنگ با استفاده از تشخیص پوست .....	۳۸

۴۲.....	۳-۴ معیارهای ارزیابی.....
۴۴.....	<b>۴ فصل چهارم: روش پیشنهادی و ارائه نتایج.....</b>
۴۵.....	۴-۱ مجموعه داده.....
۴۵.....	۴-۲ چالش‌ها.....
۴۶.....	۴-۳ کاهش ابعاد.....
۴۸.....	۴-۴ ساخت واژه‌نامه.....
۴۹.....	۴-۵ ساخت بسته‌های کلمات بصری.....
۵۰.....	۴-۶ استخراج ویژگی عمومی.....
۵۲.....	۴-۷ نتایج دسته‌بندی کننده‌های پایه.....
۶۵.....	۴-۸ دسته‌بندی تصاویر.....
۶۷.....	۴-۹ نتایج دسته‌بندی کننده‌های نهایی.....
۷۰.....	۴-۱۰ مقایسه روش پیشنهادی با کارهای پیشین.....
۷۲.....	<b>۵ فصل پنجم: نتیجه‌گیری و کارهای آینده.....</b>
۷۲.....	۵-۱ نتیجه‌گیری.....
۷۴.....	۵-۲ کارهای آتی.....
۷۶.....	<b>۶ منابع.....</b>

## فهرست شکل‌ها و نمودارها

- شکل ۱-۲: گروه‌بند شمایل انسان پیشنهادشده توسط فلک و همکاران [۴]..... ۱۰
- شکل ۲-۲: فرآیند کلی سیستم پیشنهادی وانگ و همکاران [۷]..... ۱۱
- شکل ۳-۲: تصاویر تشخیص داده‌شده در سیستم جونز و رنگ [۸]..... ۱۱
- شکل ۴-۲: خطوط صاف تشخیص داده‌شده در عکس به‌عنوان یک ویژگی [۱۶]..... ۱۴
- شکل ۵-۲: تصاویر نمونه از پنج کلاس مختلف تصاویر در تحقیق دیسلرز و همکاران [۱۷]..... ۱۴
- شکل ۱-۳: شمای کلی از یک سیستم بازیابی تصویری [۲۴]..... ۱۷
- شکل ۲-۳: ساخت واژه‌نامه بصری [۲۴]..... ۱۹
- شکل ۳-۳: توصیف یک تصویر به شکل برداری از بسامد کلمات بصری [۲۴]..... ۲۰
- شکل ۴-۳: تأثیر اندازه تصویر در تعداد نواحی تشخیص داده‌شده [۲۴]..... ۲۶
- شکل ۵-۳: تشخیص چندین ناحیه برای یک قسمت از عکس [۲۴]..... ۲۹
- شکل ۶-۳: شباهت بین دو تصویر [۳۲]..... ۳۴
- شکل ۷-۳: مشتق مرتبه دوم گاوسی و تخمین این توابع در SURF [۳۲]..... ۳۵
- شکل ۸-۳: افزایش سایز فیلتر به‌جای کاهش سایز تصویر [۳۲]..... ۳۶
- شکل ۹-۳: ویژگی‌های طبیعی تشخیص‌دهنده‌های بر پایه هسیان [۳۲]..... ۳۶
- شکل ۱۰-۳: نواحی انطباق یافته در دو تصویر با اعوجاج افاین کامل [۳۴]..... ۳۸
- شکل ۱۱-۳: تصاویر نشانه‌گذاری برای تشخیص پیکسل‌های پوست در مرحله‌ی ساخت مدل... ۴۱
- شکل ۱۲-۳: تصویر شباهت به پوست و تصاویر فیلتر شده..... ۴۱
- شکل ۱-۴: تمرکز توصیفگرها روی مناطق هم‌رنگ پوست..... ۴۷
- شکل ۲-۴: تعداد تصاویر غیراخلاقی و معمولی با نرخ پوست‌های متفاوت..... ۵۱

- شکل ۳-۴: تشخیص صورت و تعیین نرخ پوست..... ۵۲
- شکل ۴-۴: نرخ مثبت اشتباه برای روش‌های مختلف وزندهی به واژه‌نامه msr..... ۵۳
- شکل ۵-۴: نرخ دقت برای روش‌های مختلف وزندهی به واژه‌نامه msr..... ۵۴
- شکل ۶-۴: نرخ بازیافت برای روش‌های مختلف وزندهی به واژه‌نامه msr..... ۵۴
- شکل ۷-۴: معیار اف برای روش‌های مختلف وزندهی به واژه‌نامه msr..... ۵۵
- شکل ۸-۴: نرخ مثبت اشتباه و بازیافت برای روش‌های مختلف وزندهی به واژه‌نامه msr..... ۵۵
- شکل ۹-۴: نرخ مثبت اشتباه برای روش‌های مختلف وزندهی به واژه‌نامه SURF..... ۵۹
- شکل ۱۰-۴: معیار دقت برای روش‌های مختلف وزندهی به واژه‌نامه SURF..... ۵۹
- شکل ۱۱-۴: معیار بازیافت برای روش‌های مختلف وزندهی به واژه‌نامه SURF..... ۶۰
- شکل ۱۲-۴: معیار اف برای روش‌های مختلف وزندهی به واژه‌نامه SURF..... ۶۰
- شکل ۱۳-۴: نرخ مثبت اشتباه و بازیافت برای روش‌های مختلف وزندهی به واژه‌نامه SURF..... ۶۱
- شکل ۱۴-۴: نتایج دسته‌بندی کننده‌های مختلف با ویژگی‌های عمومی..... ۶۵
- شکل ۱۵-۴: ترکیب دسته‌بندی کننده‌ها بر مبنای الگوریتم رأیگیری..... ۶۶
- شکل ۱۶-۴: ترکیب دسته‌بندی کننده‌ها بر مبنای الگوریتم تولید پشته..... ۶۶
- شکل ۱۷-۴: سیستم پیشنهادی برای ترکیب دسته‌بندی کننده‌ها..... ۶۷
- شکل ۱۸-۴: مقایسه بین الگوریتم‌های مختلف دسته‌بندی تصاویر غیراخلاقی با استفاده از ویژگی‌های عمومی و محلی بر اساس معیارهای بازیافت، دقت و اف..... ۶۹
- شکل ۱۹-۴: مقایسه بین الگوریتم‌های مختلف دسته‌بندی تصاویر غیراخلاقی با استفاده از ویژگی‌های عمومی و محلی بر اساس معیار نرخ مثبت اشتباه..... ۶۹
- شکل ۲۰-۴: مقایسه باکارهای پیشین..... ۷۱

## فهرست جدول‌ها

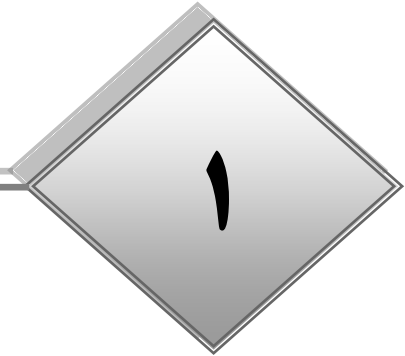
- جدول ۱-۳: وزن محلی واژه ..... ۳۱
- جدول ۲-۳: وزن دهی سراسری به واژه ..... ۳۲
- جدول ۳-۳: فاکتور نرمال‌ساز برای سند ..... ۳۳
- جدول ۴-۱: نتایج دسته‌بندی الگوریتم پایه جنگل تصادفی با روش‌های مختلف وزندهی به واژه‌نامه  
MSER ..... ۵۶
- جدول ۴-۱: نتایج دسته‌بندی الگوریتم پایه جنگل تصادفی با روش‌های مختلف وزندهی به واژه‌نامه  
SURF ..... ۶۱
- جدول ۴-۲: نتایج دسته‌بندی کننده‌های مختلف با ویژگی‌های عمومی نرخ پوست کل تصویر و  
نرخ پوست زیر صورت ..... ۶۴
- جدول ۴-۳: نتایج دسته‌بندی کننده‌های مختلف با ویژگی عمومی نرخ پوست کل تصویر ..... ۶۴
- جدول ۴-۴: نتایج دسته‌بندی کننده‌های پایه روی مجموعه ویژگی‌های عمومی و محلی ..... ۶۸
- جدول ۴-۵: نتایج دسته‌بندی کننده‌های ترکیبی ..... ۶۸
- جدول ۴-۶: مقایسه باکارهای پیشین ..... ۷۰



## چکیده

نام خانوادگی: ترابیان دهکردی	نام: علی	شماره دانشجویی: ۹۱۱۵۴۰۱
عنوان پایان نامه: فیلتر تصاویر غیراخلاقی براساس تشخیص پوست		
استاد/ اساتید راهنما: دکتر سید عنایت‌الله علوی		
استاد/ اساتید مشاور: دکتر علیرضا عصاره		
درجه تحصیلی: کارشناسی ارشد	رشته: مهندسی کامپیوتر	گرایش: هوش مصنوعی
دانشگاه: شهید چمران اهواز	دانشکده: مهندسی	گروه: کامپیوتر
تاریخ فارغ‌التحصیلی: تابستان ۱۳۹۳		تعداد صفحه: ۸۷
کلیدواژه‌ها: تشخیص تصاویر غیراخلاقی، تشخیص پوست، ویژگی‌های محلی، بسته کلمات بصری، ترکیب دسته-بندها		
<p><b>چکیده:</b> هدف از پژوهش حاضر ارائه روشی برای تشخیص و جداسازی تصاویر غیراخلاقی از تصاویر معمولی است. تشخیص تصاویر غیراخلاقی برای فیلتر مفاهیم نامناسب در اینترنت لازم و ضروری است. به‌خصوص در مواردی همچون فیلتر مفاهیم نامناسب در شبکه‌های اجتماعی و اشتراک‌گذاری تصویر، که از فیلتر کل سایت پرهیز می‌شود. تنوع در مناظر، پس‌زمینه‌ی شلوغ و نواحی غیراخلاقی کوچک در تصویر از چالش‌های این مسئله هستند. تشخیص تصاویر غیراخلاقی برپایه بسته‌های کلمات بصری یک رویکرد امیدوارکننده است زیرا می‌تواند کاستی‌های روش‌های سنتی را جبران کند. با این حال انتخاب‌های بسیاری برای ساخت کلمات بصری وجود دارد که در ایجاد موازنه بین سرعت و کارایی، حیاتی هستند. در این پژوهش روش جدیدی برای ساخت بسته‌های کلمات SURF و MSER در نواحی هم‌رنگ پوست و ترکیب آن با ویژگی‌های عمومی استخراج‌شده از نرخ پوست ارائه کرده‌ایم. نشان دادیم نتایج سیستم با ترکیب این ویژگی‌ها بهبود می‌یابد. TF-IDF روش معمول وزن‌دهی به کلمات است اما به آن بسنده نکردیم و برای یافتن بهترین روش وزن‌دهی، روش‌های مختلف وزن‌دهی محلی، سراسری و نرمال‌سازی را برای هر کدام از واژه‌نامه‌های SURF و MSER آزمایش کردیم. برای هر ویژگی بهترین الگوریتم دسته‌بندی را انتخاب کرده و در سیستم ترکیبی پیشنهادی استفاده کردیم. نتایج نشان می‌دهد سیستم ترکیبی پیشنهادی می‌تواند تصاویر غیراخلاقی را با نرخ بازیافت (مثبت صحیح) ۹۶٫۱ درصد تشخیص دهد درحالی‌که تنها نرخ مثبت اشتباه ۵٫۷ درصد را نشان می‌دهد.</p>		

## مقدمه



### ۱-۱ تعریف مسئله

با رشد سریع اینترنت، جستجو و دسترسی به فهرست پرباری از مطالب چندرسانه‌ای برای هر کاربر فراهم شده است. اگرچه اینترنت با آسان کردن اشتراک‌گذاری اطلاعات، نقش بسیار بزرگی در رشد دانش داشته است اما برخی از مطالب همچون مطالب غیراخلاقی یا به عبارتی جنسی، برای همه کاربران به‌خصوص کودکان مناسب نیست. چگونگی جلوگیری از دسترسی کاربران به این مضامین تبدیل به یک مشکل اساسی شده است. در حال حاضر محصولات تجاری و راه‌حل‌های تحقیقاتی متنوعی برای فیلترینگ<sup>۱</sup> این‌گونه مطالب وجود دارد.

بر طبق آمار بیش از ۴,۲ میلیون وب‌سایت پورنوگرافی<sup>۲</sup> در اینترنت وجود دارد (۱۲ درصد از کل وب‌سایت‌ها) و ۳۷۲ میلیون صفحه وب پورنوگرافی و ۲,۵ میلیون نامه‌ی الکترونیکی در روز (تقریباً ۸ درصد از کل نامه‌های الکترونیکی). به‌علاوه روزانه بیش از ۲۰۰۰۰ وب‌سایت پورنوگرافی به اینترنت اضافه می‌شود. [۱]

<sup>1</sup> Filtering

<sup>2</sup> pornographi

روش‌های سستی قادر به حل مشکل فیلتر مطالب غیراخلاقی نیستند. بنابراین محققان سعی کرده‌اند با استفاده از روش‌های هوشمند، به شکل خودکار صفحات غیراخلاقی را تشخیص دهند. هرچند موضوع این تحقیق فیلتر صفحات وب نیست اما تشخیص خودکار و هوشمند تصاویر غیراخلاقی قدم اول در این راه خواهد بود. به علاوه تصاویر و ویدئوهای غیراخلاقی بیشترین تأثیر مخرب را بر روی افراد جوان دارد. تجزیه و تحلیل و تشخیص تصاویر غیراخلاقی تنها راه حل این مشکل از مبحث امنیت شبکه است.

## ۲-۱ انواع روش‌های فیلترینگ اینترنت

حفاظت از کاربران فضای مجازی در مقابل هجوم مطالب غیراخلاقی یا به عبارتی «جنسی» موجود در دنیای مجازی، وظیفه‌ای انسانی و شرعی است. تحقیقات دانشمندان، از آثار سوء روحی و جسمی این مطالب بر افراد به خصوص کودکان و نوجوانان حکایت دارد. برای درک چرایی و چگونگی وجود این حجم مطالب و تصاویر غیراخلاقی و یافتن راه حل باید نگاهی به ساختار تولید و مصرف محتوا در اینترنت بیاندازیم.

در اینترنت، یک محتوا مثلاً عکس در یک محل تولید می‌شود و به سرویس‌دهنده‌ی وب<sup>۱</sup> یا همان میزبان<sup>۲</sup> سایت ارسال می‌شود. این میزبان می‌تواند در هر گوشه از جهان باشد. پس از این است که هر کاربر می‌تواند با مراجعه به میزبان، آن محتوا را دریافت کند.

برای جلوگیری از دسترسی به این نوع تصاویر، یک راه حل، جلوگیری از ساخت آن است. برای نمونه در صنعت سینما و همچنین تلویزیون قوانین و مقرراتی برای نظارت بر ساخت فیلم‌ها وجود دارد و به سبب تولیدکنندگان محدودشان قابل کنترل هستند. اما با ظهور دوربین‌های دیجیتال که

<sup>1</sup> Web server

<sup>2</sup> Host

به شکل ارزان و همه‌گیر درآمده‌اند و حتی بر روی دوربین‌های موبایل و رایانه‌ها نصب شده‌اند، نظارت بر ساخت فیلم و تصاویر از توان دولت‌ها خارج شده است.

راه‌حل بعدی نظارت بر بستر نگهداری و پخش محتوا است. در گذشته‌ی نه‌چندان دور، تنها سینما و تلویزیون و مجلات و روزنامه‌ها قادر به ارائه تصاویر به مخاطبان خود در سطح گسترده بودند که به سبب ساختار و تعداد محدود سینماها و شبکه‌ها و مجلات قابل کنترل بودند. اما با ظهور ویدئو پخش و سپس دیسک فشرده توانایی دولت و نهادهای نظارتی بر کنترل توزیع و پخش محتوا به چالش کشیده شد. در نهایت با همه‌گیر شدن اینترنت، دوران جدیدی از عمر رسانه آغاز شد به شکلی که شعار "هر کاربر یک رسانه" برآورده این دوران است.

شرط توفیق در حذف محتوای غیراخلاقی از این طریق، نظارت دقیق بر بستر توزیع یعنی اینترنت است. اما میزبان‌های وب در کشورهای مختلف با قوانین مختلف و گاه متضاد واقع شده‌اند پس نه تنها هیچ خانواده و ارگانی بلکه هیچ دولتی نمی‌تواند تمام میزبان‌ها را کنترل کند.

بنابراین تنها راه‌کار باقی‌مانده، تمرکز بر ضلع سوم یعنی کاربران است. یعنی فرض بر این است که تسلطی بر تولید، نگهداری و ارائه مطالب غیراخلاقی نداریم و تنها می‌توانیم جلوی دسترسی کاربران تحت نظارت‌مان را بگیریم. پس به دنبال راهی برای فیلترینگ پاسخ به درخواست‌های نامشروع کاربران که خواسته یا ناخواسته ارسال شده است می‌گردیم.

به شکل معمول، روش‌های فیلترینگ مطالب غیراخلاقی فعلی را می‌توان به سه دسته طبقه‌بندی

کرد<sup>[۱]</sup>:

۱. فیلترینگ بر اساس لیست سیاه <sup>۱</sup>IP

۲. فیلترینگ بر اساس محتوای متنی<sup>۲</sup>

<sup>۱</sup> IP based black list

<sup>۲</sup> Textuall content

۳. فیلترینگ بر اساس محتوای دیداری<sup>۱</sup>

### ۱-۲-۱ فیلترینگ بر اساس لیست سیاه IP

در این روش لیستی از URL سایت‌های دارای مضامین مخرب تهیه می‌شود. اگر آدرس سایتی در این لیست باشد دسترسی به آن ممنوع خواهد بود. اما محتوا در اینترنت بسیار ناپایدار است. زیرا بسیاری از وبسایت‌ها و صفحات وب زمانی نمایان می‌شوند و ممکن است زمان دیگر ناپدید گردند. در حقیقت به‌روز نگه‌داشتن لیست سیاه بسیار مشکل است. بنابراین رویکرد فیلترینگ مبتنی بر لیست سیاه ناکارآمد و غیرعملی است.

### ۲-۲-۱ فیلترینگ بر اساس محتوای متنی

این نوع فیلترینگ تلاش دارد تا با استفاده از تجزیه و تحلیل محتوای متنی، سایت‌های دارای مضامین غیراخلاقی را مسدود کند. تعدادی از کلمات کلیدی و یا عبارات توهین‌آمیز که اغلب در وبسایت‌های بزرگسالان رخ می‌دهد به شکل دستی و یا از طریق یادگیری ماشین یا فن‌های داده‌کاوی<sup>۲</sup> مشخص خواهد شد. هر کلمه یا عبارت یک صفحه وب درخواست شده، با محتوای فرهنگ کلمات کلیدی که شامل واژگان یا عبارات ممنوع است مقایسه می‌شود. اگر تعداد زیادی از کلمات یا عبارات توهین‌آمیز در یک صفحه رخ دهد آن صفحه مسدود می‌شود.

با این حال، روش فیلتر کردن مبتنی بر محتوای متنی از پدیده‌ی "بیش‌ازحد مسدود کردن"<sup>۳</sup> رنج می‌برد. پدیده‌ای که مانع از دسترسی به وبسایت‌های آموزشی از قبیل سلامتی و یا مبحث مطالعات جنسی<sup>۴</sup> می‌شود. علاوه بر این بسیاری از وبسایت‌ها با متن گنجانده شده در تصاویر از طریق فیلتر

<sup>1</sup> Visual content

<sup>2</sup> Data mining

<sup>3</sup> over-blocking

<sup>4</sup> sexology

مبثنی بر متن قابل فیلترینگ نیستند. بنابراین، بسیاری از محققان از طریق تجزیه و تحلیل تصاویر سایت‌ها نسبت به فیلتر سایت‌ها اقدام می‌کنند.

### ۱-۲-۳ فیلترینگ بر اساس محتوای دیداری

این روش فیلترینگ بر اساس تجزیه و تحلیل محتوای تصاویر نسبت به فیلتر صفحه وب یا وب‌سایت اقدام می‌کند. از مزایای این روش می‌توان به نکات زیر اشاره کرد:

با استفاده از تشخیص تصاویر غیراخلاقی می‌توان فقط یک تصویر مشخص را به جای فیلتر کل یک سایت یا حتی یک صفحه وب، فیلتر کرد. این راه‌کار به‌خصوص در مورد سایت‌های کاربر محور یعنی سایت‌هایی که کاربران محتوای آن را تولید می‌کنند بسیار راهگشا خواهد بود. به‌عنوان مثال فیلتر شبکه‌های اجتماعی (مانند فیس‌بوک و توییتر) و همچنین نرم‌افزارهای اشتراک‌گذاری عکس و متن و ویدئو که برای موبایل‌های هوشمند عرضه شده‌اند (مانند واتس‌آپ<sup>۱</sup>، وایبر<sup>۲</sup> و اینستاگرام<sup>۳</sup> و...) به معضل بزرگی برای دولت تبدیل گشته است. به شکلی که دولت نیز به‌عنوان متصدی فیلتر سراسری کشور به دنبال راه‌کاری برای خارج ساختن این سایت‌ها از فیلتر و درعین حال محافظت از کاربران در مقابل مطالب غیراخلاقی است.

همچنین تشخیص تصاویر غیراخلاقی به مدیران سایت‌ها و برنامه‌های اشتراک‌گذاری تصویر هم کمک خواهد کرد. اهمیت موضوع به‌خصوص در کشور ما از آنجا قابل درک است که ایجاد قابلیت ارسال عکس برای مدیران یک سایت ساده و کوچک و یا یک سایت درون‌سازمانی، کاری پر ریسک است. زیرا ممکن است کاربران عکس‌های غیراخلاقی را به‌جای عکس مربوط ارسال کرده و سایت موردنظر ناخواسته آن را نمایش دهد. بنابراین احتیاج به نیروی انسانی کافی برای تشخیص و فیلتر این‌گونه تصاویر ارسالی است که به‌نوبه‌ی خود موجب افزایش هزینه‌ها، کاهش راندمان، زمان‌بر شدن

<sup>1</sup> Whats app

<sup>2</sup> viber

<sup>3</sup> instagram

پاسخ به درخواست ارسال عکس، نارضایتی کاربران و درنهایت شکست خوردن این پروژه‌های کاربر محور از نمونه‌های خارجی که محدودیت‌های اخلاقی و قانونی را ندارند خواهد شد.

### ۳-۱ اهداف پایان‌نامه

تشخیص تصاویر غیراخلاقی در فیلتر مفاهیم غیراخلاقی از جمله در فیلتر سایت‌ها و صفحات وب ضروری است. اما این تمام کاربرد آن نیست. ساخت سخت‌افزارهایی همچون دوربین و تلویزیون هوشمند که با تشخیص فریم‌های غیراخلاقی از ضبط و نمایش آن‌ها خودداری کنند و یا مرورگر وب هوشمند که تصاویر غیراخلاقی را فیلتر کند یا حداقل افزونه‌ای که این قابلیت را به مرورگرهای کنونی اضافه کند از جمله کاربردهای تشخیص تصاویر غیراخلاقی خواهد بود که پیشنهاد می‌شود. همچنین دولت‌ها می‌توانند افزونه‌ای برای مرورگر وب ایجاد کرده که تنها کاربرانی که از این افزونه استفاده می‌کنند بتوانند وارد شبکه‌های اجتماعی شوند. با این روش حجم پردازش اطلاعات فیلتر سراسری هم کاهش پیدا خواهد کرد.

در این پایان‌نامه سعی شده است یک سیستم هوشمند برای تشخیص خودکار تصویر غیراخلاقی ارائه شود. در حقیقت این یک مسئله‌ی کلاس‌بندی دودویی است که یک کلاس آن تصاویر غیراخلاقی یا پورنو است و کلاس دیگر تصاویر معمولی که می‌تواند شامل هرگونه تصویری باشد. چون دسته‌بند سیستم پیشنهادی به شکل نظارت‌شده<sup>۱</sup> یا باراهنما ایجاد می‌شود، عملکرد سیستم نهایی به مجموعه داده‌ی آموزشی وابسته است. یعنی با استفاده از مجموعه آموزشی به کلاس‌بند می‌فهمانیم چه تصاویری غیراخلاقی است. چون هیچ مجموعه داده‌ی استانداردی در این زمینه وجود ندارد ما نیز همچون بسیاری کارهای مشابه، مجموعه‌ی داده‌ی تصاویر را به شکل دستی و از طریق جستجو در اینترنت ساخته‌ایم. بنابراین در اینجا نقش سلیقه پررنگ می‌شود. چون فرهنگ‌های

---

<sup>۱</sup> supervised

مختلف و افراد مختلف برداشت‌های متفاوتی از حدود تصاویر غیراخلاقی دارند ما تنها تصاویر کاملاً غیراخلاقی را در نظر گرفته‌ایم.

#### ۴-۱ روش استفاده‌شده

سعی شده است با بررسی روش‌های متنوع پردازش تصویری برای طبقه‌بندی عکس‌ها و همچنین کارهای گذشته در زمینه‌ی تشخیص تصاویر غیراخلاقی به بهترین روش در این زمینه هم از نظر دقت و هم سرعت دست پیدا کنیم. روش پیشنهادی باید بین دقت و سرعت یک توازن مناسب برقرار کند.

روش‌های سنتی عموماً مبتنی بر مفاهیم بصری عمومی سطح پایین هستند. ویژگی‌های عمومی برای یافتن ویژگی‌های کلی یک تصویر بکار می‌روند. به‌خصوص، رنگ پوست که سرعت محاسبات آن و قدرتمند بودن آن برای تشخیص تصاویر مستهجن ثابت شده است. باین حال روش سنتی انتخاب مناسبی به دو دلیل نیست: اولاً به دلیل "فاصله معنایی" [۲] موجود بین ویژگی‌های سطح پایین و غنای مفاهیم انسانی. مثلاً رنگ پوست یا تعداد خطوط صاف در تصویر برای انسان مفهوم غیراخلاقی ندارد ولی به‌عنوان یک ویژگی برای تشخیص تصاویر استفاده می‌شود. ثانیاً ویژگی‌های عمومی اغلب برای نشان دادن یک تصویر بیش‌ازحد غیرقابل انعطاف می‌باشند و موفق به شناسایی ویژگی‌های مهم تصویری نمی‌شوند [۳].

تشخیص تصاویر جنسی بر اساس بسته‌های کلمات بصری<sup>۱</sup> به علت سادگی و عملکرد خوب یک رویکرد امیدوارکننده است. هدف آن این است که الگوهای محلی رفتار مستهجن، مانند افشای اندام‌های جنسی، و نشان دادن حالت‌های غیراخلاقی را بگیرد. این الگوها برای قضاوت دقیق در

<sup>1</sup> Bag of visual words



مورد جنسی بودن یا نبودن تصاویر قاطع هستند. ویژگی‌های استخراجی از عکس در روش ما ترکیبی از بسته‌های کلمات بصری و ویژگی‌های عمومی است.

### ۵-۱ ساختار پایان‌نامه

در فصل دوم این تحقیق مروری بر کارهای پیشین صورت گرفته است. در واقع سعی شده است روش‌های متفاوت در قالب کارهای پیشین معرفی شود. در فصل سوم مبانی نظری الگوریتم‌های استفاده‌شده در این پژوهش تشریح شده است. فصل چهارم به ارائه سیستم پیشنهادی و نتایج پژوهش جاری و تجزیه و تحلیل این نتایج و همچنین مقایسه نتایج پژوهش حاضر با کارهای مشابه، پرداخته است. در فصل پنجم نتیجه‌گیری مباحث گفته‌شده صورت گرفته است و پیشنهادهایی برای کارهای آتی ارائه شده است.

تشخیص تصاویر غیراخلاقی معمولاً جزو مسائل طبقه‌بندی تصاویر، در نظر گرفته می‌شود. عموماً، ابتدا نواحی پوست با استفاده از مشخصات رنگ و بافت سگمنت می‌شوند. سپس ویژگی‌های استخراج‌شده از نواحی بخش‌بندی<sup>۱</sup> شده شامل رنگ، بافت و شکل برای جداسازی تصاویر معمولی از غیراخلاقی استفاده می‌شود. برخی از محققان از راه‌کار بازیابی تصاویر بر پایه محتوا<sup>۲</sup> برای تشخیص تصاویر غیراخلاقی استفاده کردند. در این روش، عکس‌های مشابه با عکس موردنظر در پایگاه داده حاوی تصاویر معمولی و غیراخلاقی استخراج می‌شود. اگر اکثر تصاویر بازیابی شده غیراخلاقی بودند، تصویر ورودی نیز غیراخلاقی دسته‌بندی می‌شود. در غیر این صورت تصویر موردنظر معمولی است.

روش‌های سنتی تشخیص تصاویر جنسی مبتنی بر محتوی، عمدتاً بر پایه مفاهیم بصری سطح پایین عمومی (مانند شکل، رنگ و بافت) بودند. از شکل و رنگ برای قضاوت در مورد پورنو بودن یا نبودن تصویر استفاده‌شده و چون پوست سطح صافی دارد از بافت برای کاهش مثبت اشتباه<sup>۳</sup> در

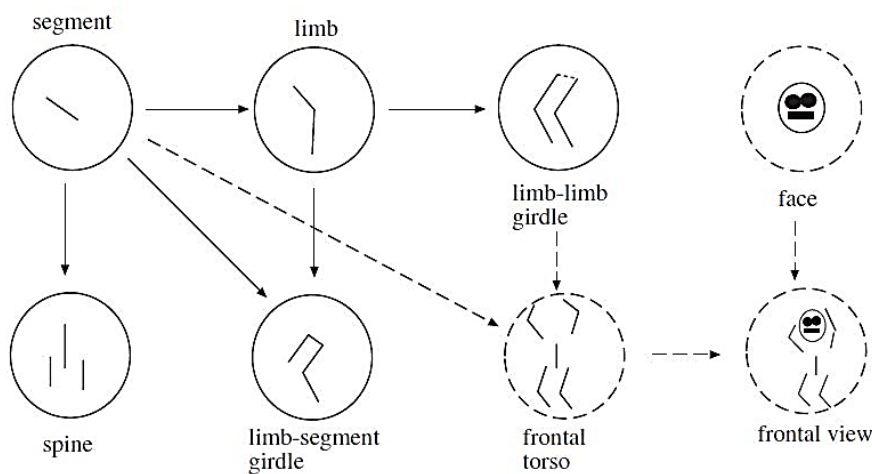
<sup>1</sup> segment

<sup>2</sup> Content based image retrieval

<sup>3</sup> False positive

هنگام تشخیص پیکسل‌های پوست استفاده می‌شود. در ادامه برخی از کارهای پیشین بررسی شده است.

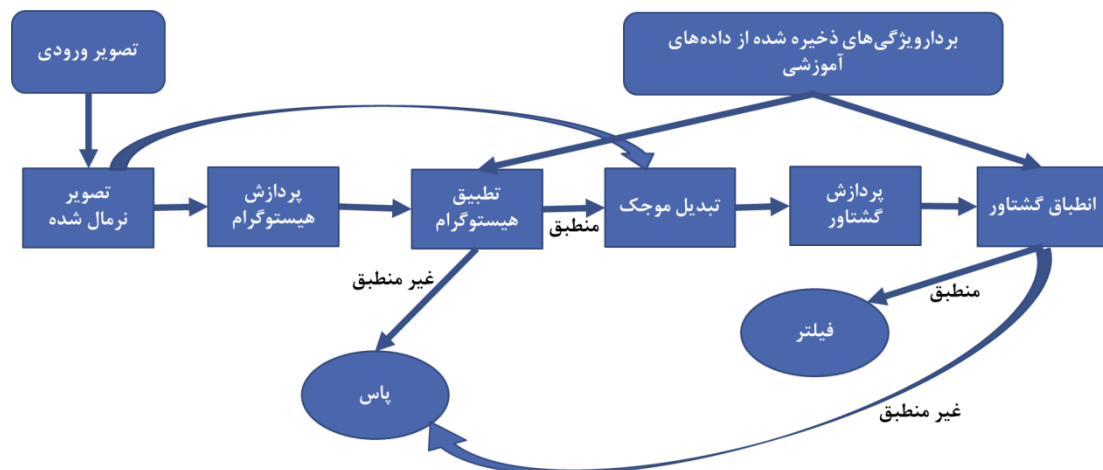
اولین تلاش در این زمینه را باید کار (فورسیت و همکاران ۱۹۹۶) [۴] [۵] [۶] دانست. آن‌ها پس از تشخیص پوست یک گروه‌بند شمایل<sup>۱</sup> انسان ساختند. از بافت و رنگ برای تشخیص نواحی پوست استفاده کردند. (شکل ۱) سپس آنالیز ژئومتریک را برای تشخیص شباهت نواحی بخش‌بندی شده با شمایل‌های انسان به کار بردند. اما زمان طولانی پردازش و دقت کم دو کاستی آن هستند.



شکل ۱-۲: گروه‌بند شمایل انسان پیشنهاد شده توسط فلک و همکاران [۴]، خط‌های نقطه‌چین شمایل‌هایی هستند که به دلیل بیش زمان بردن توسط فلک پیاده‌سازی نشدند ولی برای کامل‌تر شدن سیستم پیشنهاد شدند.

سیستم WIPE (وانگ و همکاران ۱۹۹۸) [۷] با استفاده از یک فیلتر هیستوگرام رنگ و سپس الگوریتم تطبیق شکل مبتنی بر موجک، سیستمی برای تشخیص و حذف تصاویر مستهجن ارائه کردند. (شکل ۲-۲)

<sup>۱</sup>figure



شکل ۲-۲: فرآیند کلی سیستم پیشنهادی وانگ و همکاران [۷]، این سیستم مستقیماً از تشخیص پوست بهره نمی‌برد اما فرآیند تحلیل رنگ به نوعی در پردازش هیستوگرام گنجانده شده است.

(جونز و رنگ ۲۰۰۲) [۸] یک مدل آماری رنگ برای تشخیص مناطق پوست و غیر پوست ساختند. به ازای هر عکس مجموعه‌ای از ویژگی‌ها شامل درصد پیکسل‌های پوست تشخیص داده شده، میانگین احتمال پوست بودن پیکسل‌ها، اندازه بزرگ‌ترین کامپوننت متصل پوست به پیکسل، تعداد کامپوننت‌های متصل پوست، درصد رنگ‌های بدون ورودی به هیستوگرام پوست و غیر پوست و همچنین طول و عرض تصویر. از هر دو نوع تصویر اخلاقی و غیر اخلاقی برای آموزش شبکه‌ی عصبی چندلایه برای تشخیص تصاویر غیر اخلاقی استفاده شد.



شکل ۲-۳: الف) تصاویر غیر اخلاقی صحیح تشخیص داده شده. ب) تصویر غلط غیر اخلاقی تشخیص داده شده در سیستم جونز و رنگ [۸].