

# کاربردهای هوش مصنوعی در زندگی روزمره

ارائه‌ی مطالب علمی و فنی

استاد: سرکار خانم دکتر کسائی

شماره‌ی گروه: گروه ۲ روز سه‌شنبه

اعضای گروه: ایمان محمدی، مهدی جعفری، ابوالفضل سلطانی

دانشکده‌ی مهندسی کامپیوتر

دانشگاه صنعتی شریف

بهار ۱۴۰۲

## فهرست مطالب

چکیده (ابوالفضل سلطانی).....	۲
مقدمه (ابوالفضل سلطانی).....	۲
۱. بینایی کامپیوتر تا ۲۰۱۰ (ابوالفضل سلطانی).....	۲
۱-۱. پیدایش دوربین‌ها.....	۵
۱-۲. تشخیص چهره.....	۵
۱-۳. تشخیص عمل انسان.....	۵
۱-۴. جستجوی تصویر.....	۵
۲- بینایی کامپیوتر از ۲۰۱۰ تا ۲۰۱۵ (مهدی جعفری).....	۵
۲-۱. تقسیم‌بندی معنایی.....	۵
۲-۲. پرسش و پاسخ دیداری.....	۷
۳- بینایی کامپیوتر از ۲۰۱۵ تا ۲۰۲۰ (ایمان محمدی).....	۹
۳-۱. ماشین‌های خودران.....	۱۰
۳-۲. خانه‌های هوشمند.....	۱۰
۳-۳. پزشکی.....	۱۰
۳-۴. تحصیل.....	۱۰
۳-۵. رسانه.....	Error! Bookmark not defined.
۳-۶. نظامی.....	Error! Bookmark not defined.
۳-۷. هوش جامع مصنوعی.....	Error! Bookmark not defined.

۵- مراجع (علی رحیمی اکبر و آرمین ثقفیان و شایان صالحی). **Error! Bookmark not defined.**

## چکیده (ابوالفضل سلطانی)

پردازش تصویر و بینایی ماشین از چندین دهه گذشته از حوزه‌های مهم در علوم کامپیوتر بوده است. از وقتی انسان‌ها توانستند اولین تصاویر را با دوربین‌های دیجیتالی ثبت کنند، پردازش کردن تصاویر از مسائل علوم کامپیوتر شد. تشخیص چهره افراد از اولین و پرکاربردترین مسائلی است که در این حوزه به وجود آمد و با راه‌حل‌های مختلف حل شد. با پیدایش شبکه‌های عصبی، کاربردهای پردازش تصویر به مراحل جدیدی دست یافت تا امروزه در مسائلی مانند ساخت خودروهای خودران به کمک ما بیاید. در این مقاله کاربردهای پردازش تصویر در دهه‌های مختلف و روش‌های نوین در این حوزه ذکر می‌شود.

## مقدمه (ابوالفضل سلطانی)

پردازش تصویر و بینایی ماشین از حوزه‌هایی است که تحقیق کردن در آن جذاب به نظر می‌آید. با شناخت بهتر کاربردها و روش‌های پردازش تصویر، دانشجویان کارشناسی می‌توانند بهتر در مورد تحقیق در این حوزه تصمیم‌گیری کنند. در این مقاله ابتدا به اختراع دوربین‌های دیجیتال می‌پردازیم. سپس به کاربردهای مختلف آن از جمله تشخیص چهره، جستجوی تصویر، تشخیص‌های پزشکی، خودروهای خودران، تبدیل تصاویر هوایی به نقشه و شناسایی و ردیابی اشیاء می‌پردازیم. در آخر روش‌های جدید مانند شبکه‌های عصبی و استفاده‌های آن‌ها را بررسی می‌کنیم.

## 1. بینایی کامپیوتر تا ۲۰۱۰ (مهدی جعفری)

### ۱-۱. پیدایش دوربین‌ها

در سال ۱۷۷۷ دانشمندان، تاثیر نور بر نقره کلورید را مورد توجه قرار دادند. با کمک این کشف مهم، در سال ۱۸۲۲ دانشمند فرانسوی، آقای ژوزف نیپس، نخستین عکس ماندگار را ثبت کرد. اما این عکس تا به امروز به جا نمانده است اما آقای نیپس، قدیمی‌ترین عکس ماندگار باقی‌مانده را نیز در سال ۱۸۲۵ به جا گذاشت.



تصویر ۱. از اولین عکس‌های ماندگار [۱]

از سال ۱۸۵۵ تا ۱۸۷۲، جیمز ماکسول سری تحقیقاتی در مورد درک رنگی، کوررنگی و تئوری رنگ منتشر کرد و مدال رامفورد را برای «در باب تئوری بینایی رنگ» دریافت کرد [۳]. ماکسول همچنین به استفاده از نظریه درک رنگی خود برای عکاسی رنگی علاقه‌مند بود: اگر مجموع هر سه نور بتواند هر رنگ قابل درک را بازتولید کند، می‌توان عکس‌های رنگی را با مجموعه‌ای از سه فیلتر رنگی تولید کرد. ماکسول در طول مقاله خود در سال ۱۸۵۵ پیشنهاد کرد که اگر سه عکس سیاه و سفید از یک صحنه از طریق فیلترهای قرمز، سبز و آبی گرفته شود و نسخه چاپ‌شده و شفاف تصاویر با استفاده از سه نورافکن مجهز به فیلترهای مشابه نمایش داده شود [۴].

ایستمن کداک در سال ۱۹۷۴، وظیفه دیجیتالی کردن تصاویر را به استیون ساسون، مهندس الکترونیک، واگذار کرد. ساسون و تیمش با استفاده از کنار هم گذاشتن قسمت‌های مختلف دوربین Kodak XL55 و همچنین ساخت یک واحد پخش برای تبدیل داده‌های نوار به تصاویر دیجیتالی یک دوربین ساختند که می‌توان تصاویر آن را روی صفحه تلویزیون نمایش داد. در ۱۹۷۵، ساسون اولین عکس خود را با استفاده از دوربین گرفت، پرتره ای از یک تکنسین آزمایشگاه کداک [۵].



تصویر ۲. تقسیم‌بندی معنایی یک ظرف حاوی چند میوه

با توجه به توانایی ذخیره تصاویر در کامپیوترها، حوزه پردازش تصاویر ایجاد شد. تصاویر در کامپیوترها هم می‌تواند با استفاده از نمایش برداری [۶]، و هم با استفاده از نمایش پیکسلی ذخیره شوند. برتری نمایش برداری این است که با بزرگ کردن تصویر، کیفیت نمایش آن تغییری نمی‌کند. در نمایش پیکسلی یک جدول از اعداد بین ۰ و ۱ یا اعداد بین ۰ تا ۲۵۵ ذخیره می‌شود و در صورت رنگی بودن تصویر، در سه جدول متمایز نگهداری می‌شود. با تعریف انتزاعی جدولی، محتوای سازنده جدول می‌تواند نور مرئی (عکس‌های عادی)، اشعه ایکس، و یا اشعه مادون قرمز باشد.

## ۲-۱. تشخیص چهره

در سال ۱۹۶۰، پروژه‌ای به نام «انسان-ماشین» در تشخیص چهره پیشگام شد. دلیل این نام‌گذاری این است که ابتدا یک انسان با استفاده از رابط گرافیکی، نقاط مهم صورت (مانند مرکز مردمک، گوشه‌های داخلی و خارجی چشم و غیره) را در یک تصویر مشخص می‌کرد. در سال ۱۹۷۰ سامانه جدیدی ایجاد شد که کامپیوتر می‌توانست عکس‌های ساده با تنها یک صورت و با محیط اطراف کنترل‌شده را بدون کمک یک انسان، تشخیص بدهد. در سال ۱۹۹۳ شرکتی در آمریکا تاسیس شد که فناوری تشخیص چهره خود را در دفترهای وسایل نقلیه موتوری به کار گرفتند تا از گرفتن چند گواهینامه توسط یک شخص و با نام‌های مختلف، جلوگیری کنند [۷]. در سال ۱۹۹۹، ایالت مینه‌سوتا آمریکا، سامانه تشخیص چهره‌ای در اختیار پلیس، قاضیان، و افسران دادگاه قرار داد که به آنها قابلیت ردیابی مجرمین در سراسر ایالت می‌داد [۸]. در روش‌های نوین تشخیص چهره، از روش تحلیل مولفه‌های اصلی و به دست آوردن ویژه‌چهره‌ها، یا یادگیری عمیق بهره گرفته می‌شود.

از کاربردهای روزمره تشخیص چهره، می‌توان به کاربرد در شبکه‌های اجتماعی اشاره کرد که با استفاده از یک فیلتر، چهره کاربران در فیلم‌ها یا تماس‌های تصویری تغییر می‌کند [۹]. همچنین تشخیص چهره در سامانه‌های امنیتی برای ورود به سامانه و یا ردیابی افراد با دوربین‌های محلی استفاده شود [۱۰].

### ۳-۱. تشخیص عمل انسان

با استفاده از تشخیص عمل انسان‌ها از روی تصاویر، راه جدیدی برای ارتباط انسان با کامپیوتر به مانند موشواره ایجاد می‌شود. این راه جدید، می‌تواند مانند بقیه ابزارهای مشابه خود، سبک زندگی انسان‌ها را تغییر دهد. همچنین این روش، دنیای جدیدی از بازی‌های کامپیوتری را معرفی می‌کند.

از کاربردهای دیگر تشخیص عمل انسان، می‌توان به تشخیص اعمال مجرمانه از دوربین‌های امنیتی، و تحلیل آماری ورزش‌هایی مثل فوتبال اشاره کرد.

### ۴-۱. جستجوی تصویر

جستجوی تصاویر می‌تواند به صورت متنی و یا تصویری باشد. در جستجوی متنی، کاربر یک متن را به عنوان پرسش وارد می‌کند و انتظار دارد عکس‌های مرتبط با آن متن را بیابد. در جستجوی تصویری، کاربر پرسش خود را به صورت یک تصویر وارد می‌کند و انتظار دارد تصاویر مشابه با آن را ببیند. از معروف‌ترین سامانه‌های جستجوی تصویر می‌توان به گوگل لنز اشاره کرد [۱۱].

## ۲- بینایی کامپیوتر از ۲۰۱۰ تا ۲۰۱۵ (ابوالفضل سلطانی)

### ۱-۲. تقسیم‌بندی معنایی

به دسته‌بندی کردن عکس به دسته‌های مختلف که هر دسته معنای خودش را داشته باشد، تقسیم‌بندی معنایی می‌گویند. در تقسیم‌بندی معنایی به هر پیکسل یک برچسب داده می‌شود که معنای آن را نشان می‌دهد. این مسئله که یکی از پرکاربردترین مسائل دنیای پردازش تصویر است، از هزاره قبلی وجود داشته ولی از سال ۲۰۱۰ با ظهور شبکه‌های عصبی بهتر حل شده و اهمیت بیش‌تری کسب کرده است. کاربردهای این مسئله زیاد است و ۴ مورد از آن را بررسی می‌کنیم.



تصویر ۲. تقسیم‌بندی معنایی یک ظرف حاوی چند میوه

خودروهای خودران یکی از بحث‌های داغ چند سال گذشته بوده است. شرکت‌های خودروسازی بزرگ در تلاش هستند با استفاده از هوش‌های مصنوعی، خودروهایی بسازند که بدون راننده‌ی انسانی بتوانند در خیابان‌ها حرکت کنند. رانندگی به آگاهی محیطی زیادی نیاز دارد. این خودروها در هر لحظه تصاویر زیادی از اطراف خود می‌گیرند و باید آن‌ها را پردازش کند تا بتواند با تصمیم‌گیری کند.

تقسیم‌بندی معنایی در پردازش این تصاویر استفاده می‌شود. با استفاده از تقسیم‌بندی معنایی، مفهوم محیط اطراف خودرو مشخص می‌شود. همان‌طور که در این تصویر می‌بینید، خودرو در حال حرکت است و در هر لحظه، تمام ماشین‌ها، جاده‌ها، انسان‌ها، موانع، و درخت‌ها مشخص هستند و بر اساس موقعیت آن‌ها تصمیم به حرکت یا چرخش می‌کند.



تصویر ۴. تقسیم‌بندی معنایی محیط اطراف یک خودرو خودران

تقسیم‌بندی معنایی در این مورد باید دقت زیادی داشته باشد، زیرا هر اشتباه آن هزینه‌های زیادی به جان و مال انسان‌ها وارد می‌کند. مجموعه داده مناظر شهری یک مجموعه داده است که از مناظر شهری تصویر واقعی و تقسیم‌بندی معنایی آن را با ۸ برچسب سطح صاف، انسان‌ها، ماشین‌ها، سازه‌ها، اشیاء، طبیعت، آسمان و خالی تشکیل شده است.

در این نمودار، دقت مدل‌های آموزش دیده روی این مجموعه داده را می‌توانید ببینید. محور X آن زمان و محور Y آن دقت مدل‌ها با معیار Mean IOU است که میانگین اشتراک به روی اجتماع پیکسل‌های درست است. در سال‌های ۲۰۱۵ با روش DeepLab دقت حدود ۶۲ درصد داشتیم و با گذر زمان پیشرفت‌های زیادی حاصل شده تا در سال اخیر با روش InternImage-H به دقتی حدود ۸۸ درصدی برسیم که دقت قابل قبولی است.



تصویر ۵. پیشرفت مدل‌های مختلف در مجموعه داده مناظر شهری طی گذر زمان

این روزها استفاده از نقشه در مسیریابی‌ها کمک زیادی به همه مردم کرده است. این نقشه‌ها چه شکلی ساخته می‌شوند؟ یک عکس هوایی از سطح زمین را در نظر بگیرید. با تقسیم‌بندی معنایی آن و شناسایی خانه‌ها، جاده‌ها آن تبدیل به نقشه می‌شود.

یک کاربرد خیلی جالب دیگر از این موضوع در زمان‌هایی است که زلزله آمده است. با استفاده از پهپادها تصاویر هوایی می‌گیرند و با تقسیم‌بندی معنایی انسان‌هایی که در آوار گرفتار شده‌اند را پیدا می‌کنند و به کمک آن‌ها می‌روند.

## ۲-۲. پرسش و پاسخ دیداری

پرسش و پاسخ دیداری ترکیب دو حوزه پردازش تصویر و پردازش زبان طبیعی است. در این مسئله یک تصویر و یک سوال در مورد تصویر پرسیده می‌شود و به دنبال دریافت جواب هستیم. به این مثال‌ها توجه کنید.





تصویر ۶. نمونه پرسش در مسئله پرسش و پاسخ دیداری

کمک به افراد نابینا از اهداف چند سال اخیر کاربردهای پرسش و پاسخ دیداری است. دلیل اصلی آن جواب دادن به پرسش‌های روزمره افراد است که می‌تواند به افراد نابینا کمک بساری بکند. در ۱۰ سال اخیر پیشرفت‌های عظیمی در این زمینه اتفاق افتاده و سرمایه‌گذاران زیادی روی این حوزه سرمایه‌گذاری کرده‌اند [۱۲].

### ۳- بینایی کامپیوتر از ۲۰۱۰ تا ۲۰۱۵ (ابوالفضل سلطانی)

#### ۳-۱. شناسایی و ردیابی اشیا

شناسایی و ردیابی اشیا از کاربردهای اساسی پردازش تصویر و بینایی کامپیوتر است. از طریق الگوریتم‌هایی مثل YOLO، می‌توان اشیا مختلفی را در تصاویر یا ویدیوها شناسایی و ردیابی کرد. این کاربردها شامل تشخیص چهره، شمارش افراد در تصویر و تشخیص هویت از روی عنبیه چشم است. همچنین، استخراج شیء از تصویر یا ویدیو، تشخیص خنده و کاربردهای رباتیک نیز از این موضوعات است. در نهایت، این تکنولوژی کمک می‌کند تا ما مدل‌های دقیقی را برای تشخیص اشیا بسازیم و در پردازش تصویر به دقت بیشتری برسیم.

## ۲-۳. صنایع

به علاوه، شناسایی و ردیابی اشیا در بسیاری از صنایع مانند خودروسازی، کنترل ترافیک، امنیت و نظارت، هوشمند سازی خانه‌ها و اماکن عمومی، پزشکی و حتی تبلیغات کاربرد دارد. به عنوان مثال، در صنعت خودروسازی، سیستم‌های خودروهای خودران برای شناسایی موانع و نشانه‌های راهنمایی استفاده می‌کنند. در حوزه پزشکی، تصاویر پزشکی می‌توانند برای شناسایی و تشخیص بیماری‌ها مانند سرطان استفاده شوند. همچنین، در حوزه تبلیغات، تشخیص چهره می‌تواند برای تهیه تبلیغات شخصی‌سازی شده استفاده شود.

## ۳-۳. رباتیک

در حوزه رباتیک نیز، شناسایی و ردیابی اشیا ضروری است. ربات‌ها می‌توانند از این تکنیک‌ها استفاده کنند تا با محیط اطراف خود تعامل داشته باشند، مانند جمع‌آوری اشیا یا اجرای وظایف خاص. در نهایت، شناسایی و ردیابی اشیا یکی از مهم‌ترین بخش‌های پردازش تصویر و بینایی کامپیوتر است که به طور گسترده‌ای در صنایع مختلف کاربرد دارد.

## ۶-۲- بازی‌های ویدیویی و واقعیت مجازی

به علاوه، در حوزه بازی‌های ویدیویی و واقعیت مجازی، پردازش تصویر و شناسایی اشیا می‌تواند تجربه کاربران را بهبود ببخشد. با تشخیص اشیا و گردش در محیط‌های واقعیت مجازی، بازی‌ها می‌توانند واکنش‌های بیشتری را از کاربران خود دریافت کنند و تجربه‌های واقع‌بینانه‌تری را ارائه دهند.

## ۳- بینایی کامپیوتر از ۲۰۲۰ به بعد (ایمان محمدی)

## ۲-۳. مدل‌های تولیدکننده برای ساخت تصاویر

مدل‌های تولیدکننده یا generative models، ابزار قدرتمندی در پردازش تصویر هستند که قادرند تصاویر جدیدی را ایجاد کنند. این مدل‌ها شامل شبکه‌های مولد تخصصی (GANs) و مدل‌های انتشاری هستند. GANs.

که توسط ایان گودفلاو معرفی شدند، از دو شبکه عصبی، یکی به عنوان مولد و دیگری به عنوان تمییزدهنده، استفاده می‌کنند که با یکدیگر رقابت می‌کنند.

## ۱-۳- مدل‌های انتشاری

مدل‌های انتشاری، یک روش دیگر برای تولید تصاویر هستند که بر اساس یک فرآیند تصادفی عمل می‌کنند و DALL-E 2 و Stable Diffusion و Midjourney، نمونه‌هایی از تکنیک‌های جدید در این حوزه هستند که به بهبود کیفیت تصاویر تولید شده و کاهش مشکلات موجود در مدل‌های قدیمی‌تر کمک کرده‌اند. سعی می‌کنند تا توزیع واقعی داده‌ها را بازتولید کنند.

## ۲-۳- نمونه ۲ DALL-E

DALL-E 2 از تکنیک‌های GAN استفاده می‌کند تا تصاویر خلاقانه و جدیدی را از روی مجموعه‌ای از داده‌های آموزشی تولید کند. این مدل، با استفاده از تکنیک‌های یادگیری عمیق، قادر است تصاویری را تولید کند که خلاقانه، منحصر به فرد و اغلب دقیقاً با توضیحات متنی مطابقت دارند.

## ۳-۳- نمونه STABLE DIFFUSION

Stable Diffusion نوعی از مدل انتشاری است که یک فرآیند تصادفی را برای تولید تصاویر جدید از داده‌های آموزشی مدل‌سازی می‌کند. این مدل، می‌تواند تصاویری با کیفیت بالا و با جزئیات زیاد تولید کند، که بسیار نزدیک به توزیع واقعی داده‌ها هستند.

## ۴-۳- نمونه MIDJOURNEY

Midjourney، یک تکنیک جدید در حوزه GANs است که سعی دارد برخی از مشکلات مربوط به تولید تصاویر با کیفیت پایین و ناهمگنی در تصاویر تولید شده توسط مدل‌های GAN قدیمی‌تر را حل کند. این تکنیک، با تغییر در روش آموزش مدل، سعی دارد تا تصاویر با کیفیت بالاتر و یکنواخت‌تری تولید کند.

این مدل‌های تولیدکننده در حوزه‌های مختلف کاربرد دارند. از جمله کاربردهای آنها می‌توان به تولید هنر دیجیتال، کمک به طراحان گرافیکی، تولید محتوای سرگرمی مانند بازی‌های ویدیویی، و حتی در پزشکی برای تولید تصاویر آناتومیکی اشاره کرد.

به علاوه، مدل‌های تولیدکننده می‌توانند در تولید داده‌های آموزشی مصنوعی کمک کنند، که برای آموزش مدل‌های یادگیری عمیق دیگری مفید است. برای مثال، در حوزه‌ی تشخیص بیماری‌های پوستی، می‌توان از مدل‌های GAN برای تولید تصاویر پوست مصنوعی با نشانگان مختلف بیماری استفاده کرد.

مدل‌های تولیدکننده نیز به عنوان ابزاری برای تحلیل و فهم بهتر داده‌ها مورد استفاده قرار می‌گیرند. با تحلیل تصاویر تولید شده توسط این مدل‌ها، ما می‌توانیم فهم بهتری از ویژگی‌ها و الگوهای موجود در داده‌های واقعی پیدا کنیم.

با این حال، استفاده از مدل‌های تولیدکننده نیز چالش‌های خود را دارد. برای نمونه، از جمله مشکلات مهم در حوزه GAN، تولید تصاویر با کیفیت پایین یا ناهمگنی است. همچنین، این مدل‌ها می‌توانند برای اهداف غیراخلاقی مانند تولید تصاویر جعلی یا سوء استفاده از اطلاعات شخصی استفاده شوند.

با این حال، با پیشرفت‌های اخیر در این زمینه، امیدواریم که بتوانیم از این تکنیک‌ها به طور ایمن و موثرتری استفاده کنیم. برای مثال، مدل‌های جدیدی مانند **Stable Diffusion** و **Midjourney**، قادر به تولید تصاویر با کیفیت بالاتر هستند، و تکنیک‌های جدید برای حفظ حریم خصوصی و امنیت داده‌ها در حال توسعه هستند.

در نهایت، مدل‌های تولیدکننده در پردازش تصویر، ابزارهای قدرتمندی هستند که در سال‌های اخیر پیشرفت‌های بسیاری داشته‌اند. با استفاده از این مدل‌ها، ما قادر به تولید تصاویر جدید، فهم بهتر داده‌های موجود و حتی تولید داده‌های آموزشی مصنوعی هستیم. با پیشرفت‌های بیشتر، ما امیدواریم که بتوانیم از این تکنیک‌ها به طور کارآمدتر و ایمن‌تری استفاده کنیم.

## ۴- نتیجه‌گیری (ایمان محمدی)

از سال ۲۰۰۰ تاکنون، پردازش تصویر و بینایی کامپیوتر پیشرفت‌های چشمگیری را شاهد بوده‌اند. با پیشرفت تکنولوژی و افزایش قدرت پردازش کامپیوترها، امکانات جدیدی در این حوزه ممکن شده‌اند.

شناسایی و ردیابی اشیاء، که از الگوریتم‌های پیچیده‌ای مانند YOLO استفاده می‌کند، به ما اجازه می‌دهد تا اشیاء مختلف را در تصاویر و ویدیوها تشخیص دهیم. این تکنیک‌ها در حوزه‌های متنوعی مانند رباتیک، امنیت، و حتی بازی‌های ویدیویی کاربرد دارند.

همچنین، با پیشرفت مدل‌های تولیدکننده مانند شبکه‌های مولد تخصصی (GANs) و مدل‌های انتشاری، قادر شده‌ایم تا تصاویر جدید و خلاقانه را ایجاد کنیم. این مدل‌ها در حوزه‌هایی مانند هنر دیجیتال، تولید محتوای سرگرمی، و حتی پزشکی کاربرد دارند.

با این حال، با این پیشرفت‌ها، چالش‌های جدیدی نیز به وجود آمده‌اند، از جمله نیاز به حفظ حریم خصوصی و امنیت داده‌ها، و مشکلات مربوط به کیفیت و یکنواختی تصاویر تولید شده. با این حال، با پیشرفت‌های اخیر در این زمینه، امیدواریم که بتوانیم از این تکنیک‌ها به طور ایمن و موثرتری استفاده کنیم.

در نهایت، پیشرفت‌های اخیر در پردازش تصویر و بینایی کامپیوتر، امکانات جدیدی را در حوزه‌های متنوعی از جمله صنعت، پزشکی، امنیت، و سرگرمی ارائه می‌دهد. این تکنیک‌ها با دادن توانایی به کامپیوترها برای درک و تفسیر دنیای اطراف، نقشی بسیار مهم در پیشبرد فناوری دیجیتال بازی کرده‌اند.

با توجه به رشد سریع در این زمینه، انتظار می‌رود که در سال‌های آینده شاهد پیشرفت‌های بیشتری در حوزه پردازش تصویر و بینایی کامپیوتر باشیم. این پیشرفت‌ها می‌تواند شامل بهبود الگوریتم‌های تشخیص و ردیابی اشیاء، تولید تصاویر با کیفیت بالاتر، و حل مسائل امنیتی و حفظ حریم خصوصی باشد.

در نهایت، پردازش تصویر و بینایی کامپیوتر، با قدرت تبدیل تصاویر و ویدیوها به اطلاعات قابل فهم برای کامپیوتر، یکی از ابزارهای قدرتمند و تاثیرگذار در عصر دیجیتال ما است. با استفاده از این تکنیک‌ها، ما قادر خواهیم بود به سمت یک جامعه هوشمندتر، امن‌تر و بهتر حرکت کنیم.

## ۵- مراجع

1. Gustavson, Todd (2009). *Camera: a history of photography from daguerreotype to digital*. New York: Sterling Publishing Co., Inc. [ISBN 978-1-4027-5656-6](#).

2. Gernsheim, Helmut (1986). *A Concise History of Photography* (3 ed.). Mineola, New York: Dover Publications, Inc. [ISBN 978-0-486-25128-8](#).
3. Johnson, Kevin (May 2012). "[Colour Vision](#)". University of St Andrews. [Archived](#) from the original on 11 November 2012. Retrieved 20 May 2013.
4. Maxwell, James Clerk (1855). "[Experiments on Colour, as Perceived by the Eye, with Remarks on Colour-Blindness](#)". *Transactions of the Royal Society of Edinburgh*. **21** (2): 275–298. [doi:10.1017/S0080456800032117](#). [S2CID 123930770](#). [Archived](#) from the original on 1 August 2020. Retrieved 10 March 2020. (This thought-experiment is described on pages 283–284. The short-wavelength filter is specified as "violet", but during the 19th century "violet" could be used to describe a deep violet-blue such as the colour of cobalt glass.)
5. [First Digital Camera History](#)
6. [Vector Graphics](#)
7. Gates, Kelly (2011). *Our Biometric Future: Facial Recognition Technology and the Culture of Surveillance*. NYU Press. p. 52. [ISBN 9780814732090](#).
8. Gates, Kelly (2011). *Our Biometric Future: Facial Recognition Technology and the Culture of Surveillance*. NYU Press. p. 54. [ISBN 9780814732090](#).
9. Shontell, Alyson (September 15, 2015). "[Snapchat buys Lookser, a 2-year-old startup that lets you Photoshop your face while you video chat](#)". *Business Insider Singapore*. Retrieved April 9, 2018.
10. "[Chinese police are using smart glasses to identify potential suspects](#)". [TechCrunch](#). February 8, 2018. Retrieved December 3, 2020.
11. [Google Lens](#): Search what you see
12. Visual Question Answering: which investigated applications? Silvio Barra, Carmen Bisogni, Maria De Marsico, Stefano Ricciardi