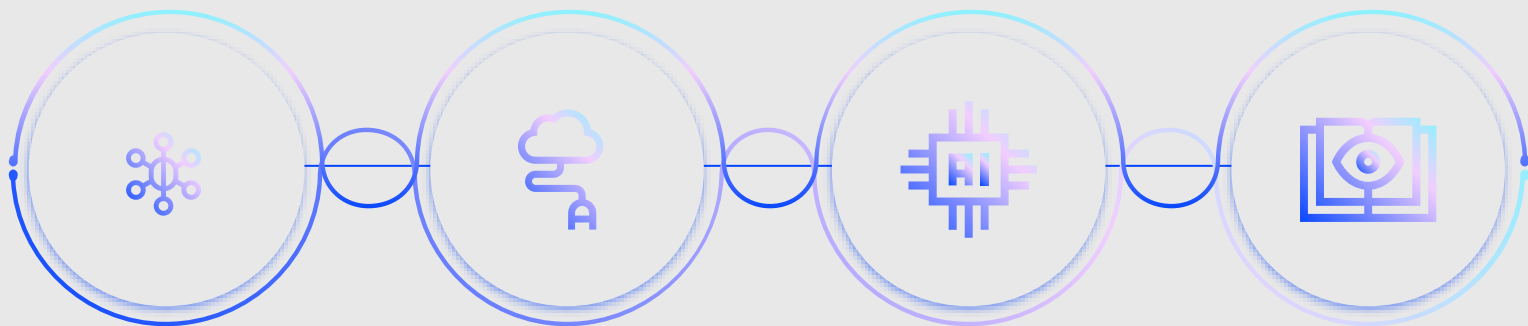


# کاربردهای پردازش تصویر و بینایی کامپیوتر

ابوالفضل سلطانی، ایمان محمدی، مهدی جعفری

پاییز ۱۴۰۲

# روند ارائه



۲۰۱۶ تا امروز

۲۰۱۱ تا ۲۰۱۵

تا ۲۰۱۰

مقدمه

# فهرست

۱. مقدمه

۲. تشخیص چهره

۳. نویسه‌خوان نوری

۴. تشخیص عمل انسان

۵. جستجوی تصاویر

۶. تقسیم‌بندی معنایی

۷. پرسش و پاسخ دیداری

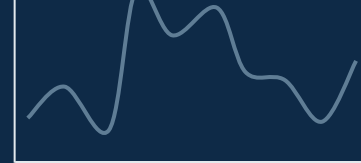
۸. شناسایی و ردیابی اشیا

۹. مدل‌های تولیدکننده برای  
ساخت تصویر

۱۰. نوشتن شرح و  
داستان‌سرایی برای تصویر

۱۱. تجزیه و تحلیل تصاویر  
پزشکی

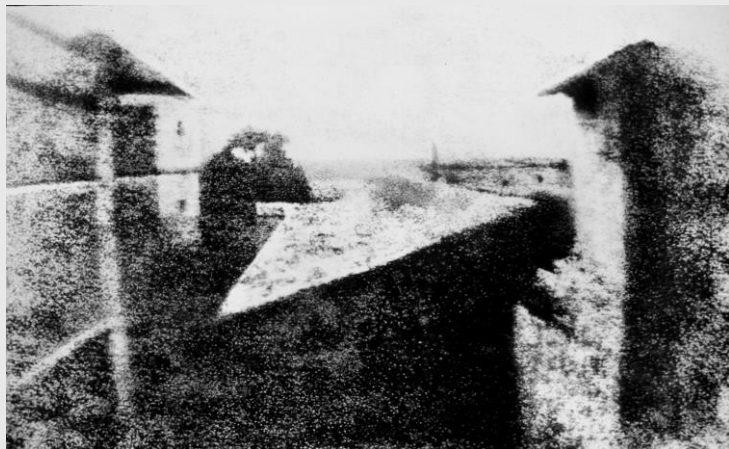
۱۲. جمع‌بندی



# مقدمه



# شروع عکاسی



از اولین عکس‌های ماندگار، ۱۸۲۶

- تاثیر نور بر نقره کلورید (Silver chloride)
- نخستین عکس ماندگار در ۱۸۲۲

# شروع عکاسی



اولین عکس رنگی، ۱۸۶۱

- تئوری سه رنگ
- جیمز ماکسول
- سه عکس سیاه و سفید با فیلتر

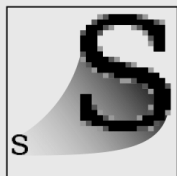
# عکاسی دیجیتال



- آقای گُذک در ۱۹۷۵
- ۳/۶ کیلوگرم
- عکس‌های مربعی با طول ۱۰۰ پیکسل



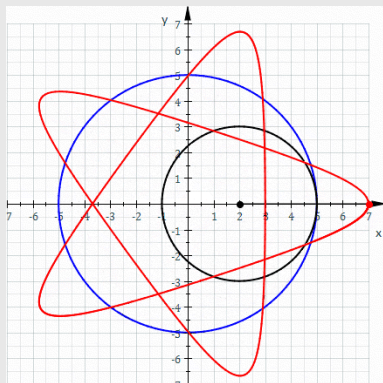
# نمایش بُرداری



Raster  
GIF, JPEG, PNG



Vector  
SVG



- تعریف شکل‌های هندسی
- خط، نقطه، و منحنی



# نمایش پیکسلی

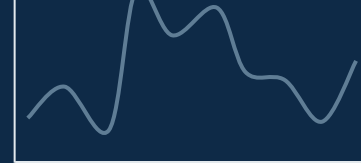
0	2	15	0	0	11	10	0	0	0	0	9	0	0	0
0	0	0	4	60	157	236	255	255	177	95	61	32	0	0
0	10	16	119	238	255	244	245	243	250	249	255	222	103	10
0	14	170	255	255	244	254	255	253	245	255	249	253	251	124
2	55	255	228	255	251	254	211	141	116	122	215	251	238	255
13	217	243	255	155	33	226	52	2	0	10	13	232	255	255
16	229	252	254	49	12	0	0	7	7	0	70	237	252	235
6	141	245	255	212	25	11	9	3	0	115	236	243	255	137
0	87	252	250	248	215	60	0	1	121	252	255	248	144	6
0	13	115	255	255	245	255	182	181	248	252	242	208	36	0
1	0	5	117	251	255	241	255	247	255	241	162	17	0	7
0	0	0	4	55	251	255	246	254	253	255	120	11	0	1
0	0	4	97	255	255	255	248	252	255	244	255	182	10	0
0	22	206	252	246	251	241	100	24	113	255	245	255	194	9
0	111	255	242	255	155	24	0	0	6	39	255	232	230	56
0	218	251	250	137	7	11	0	0	0	2	62	255	250	125
0	173	255	255	101	9	20	0	13	3	13	182	251	245	61
0	107	251	241	255	230	98	55	19	118	217	248	253	255	52
0	18	146	250	255	247	255	255	255	249	255	240	255	120	0
0	0	23	113	215	255	250	248	255	255	248	248	116	14	12
0	0	6	1	0	52	153	233	255	252	147	37	0	0	4
0	0	5	5	0	0	0	0	14	1	0	6	6	0	0

- یک یا سه لایه
- ماتریسی از اعداد
- بین ۰ و ۲۵۵

# تصویر چیست؟



- نور مرئی
- اشعه ایکس
- مادون قرمز
- هر ماتریسی از اعداد



# تشخیص چهره

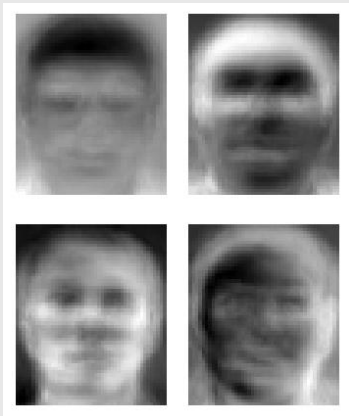
# Facial Recognition



# روش‌های ابتدایی

- ورودی: فاصله بین نقاط مهم صورت
- عکس از صورت، در محیط کنترل‌شده (مثل عکس پرسنلی)
- جلوگیری از داشتن چند گواهینامه در آمریکا، ۱۹۹۳
- سامانه قضایی آمریکا، ۱۹۹۹

# روش‌های نوین



- تشخیص نقاط مهم صورت
- به دست آوردن ویژه‌چهره‌ها با تحلیل مؤلفه‌های اصلی
- یادگیری عمیق

چند ویژه‌چهره از آزمایشگاه AT&T دانشگاه کمبریج

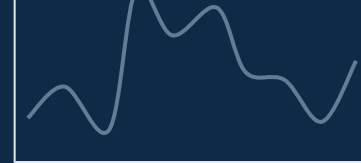
# کاربردها

۱. شبکه‌های اجتماعی

۲. سامانه‌های امنیتی



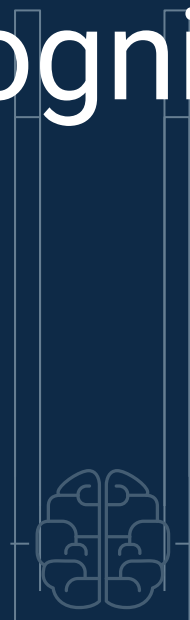
دروازه‌های ایستگاه قطار در چین، با  
استفاده از تشخیص چهره



# نویسه خوان نوری

# Optical Character

# Recognition



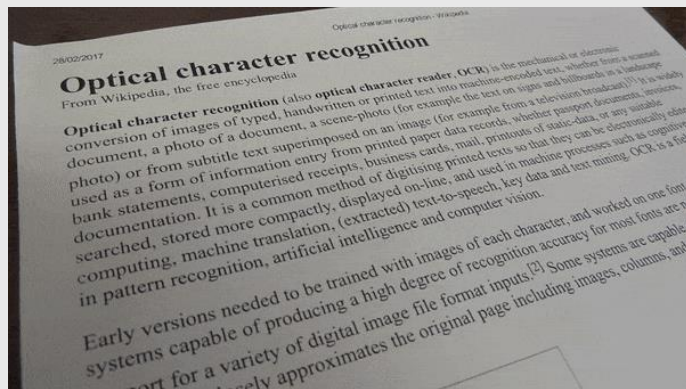


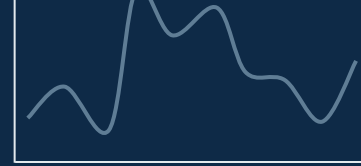
# کاربردها

۱. متن به گویش

۲. جست و جو در اسناد

۳. سامانه تشخیص پلاک





# تشخیص عمل انسان

## Human Action Recognition



# تشخیص عمل انسانی



سنسور ویژه دستگاه XBOX One

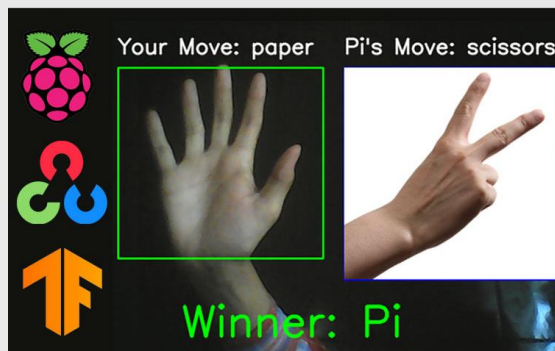
- راهی برای ارتباط انسان و کامپیوتر
- تغییر سبک زندگی
- سبک جدیدی از بازی‌های کامپیوتری

# کاربردها

۱. دوربین‌های امنیتی

۲. سرگرمی

۳. تحلیل ورزش‌ها



بازی سنگ، کاغذ، قیچی با حریف کامپیوتر،  
به صورت تصویری

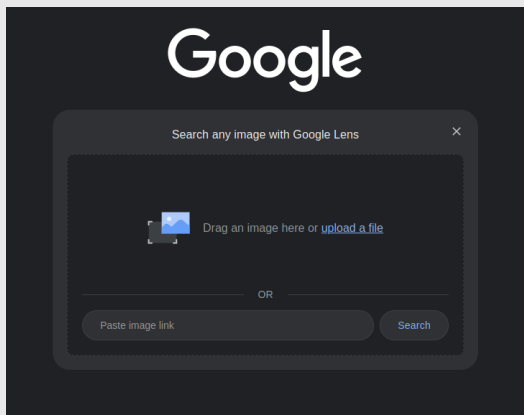


# جستجوی تصویر

# Image Retrieval



# جستجوی تصویر



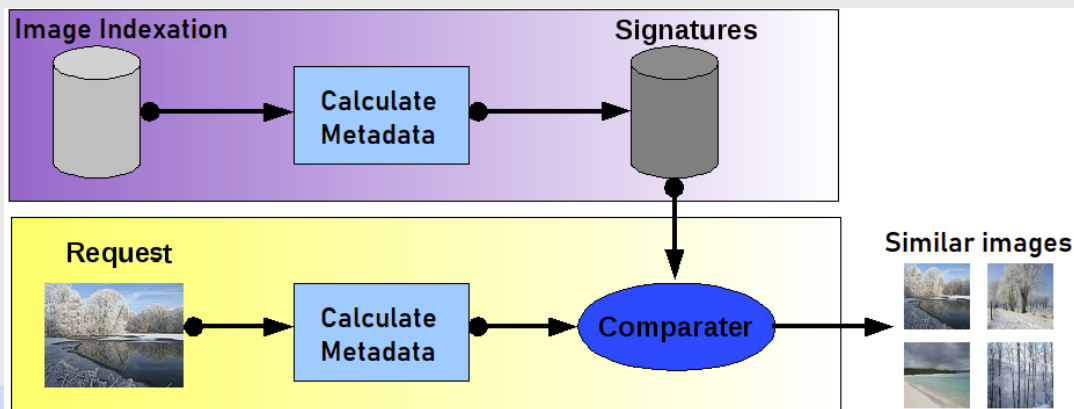
گوگل لنز، سامانه جستجوی تصویر در وب

- جستجوی متنی
- جستجوی تصویری

# کاربردها

۱. مجموعه‌های هنری

۲. تشخیص اشیا





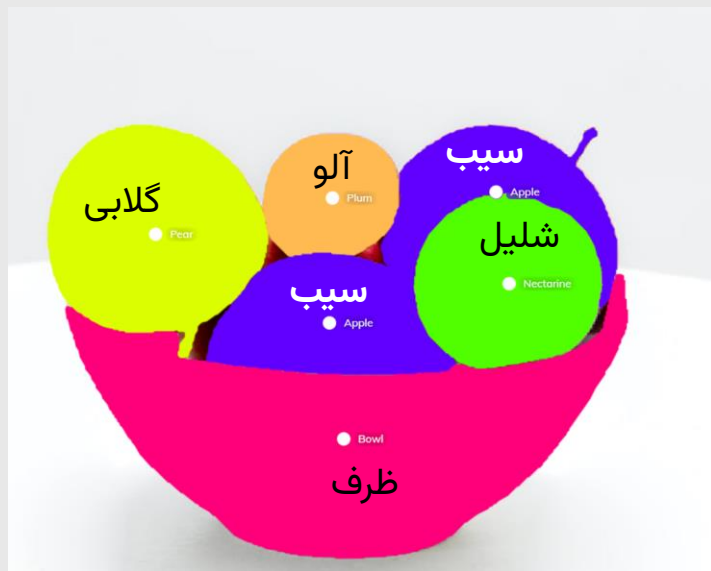


تقسیم‌بندی معنایی

# Semantic Segmentation



# تقسیم‌بندی معنایی



- تقسیم‌بندی تصویر بر اساس معنا
- برچسب زدن به نقاط مختلف تصویر

# کاربردهای تقسیم‌بندی معنایی

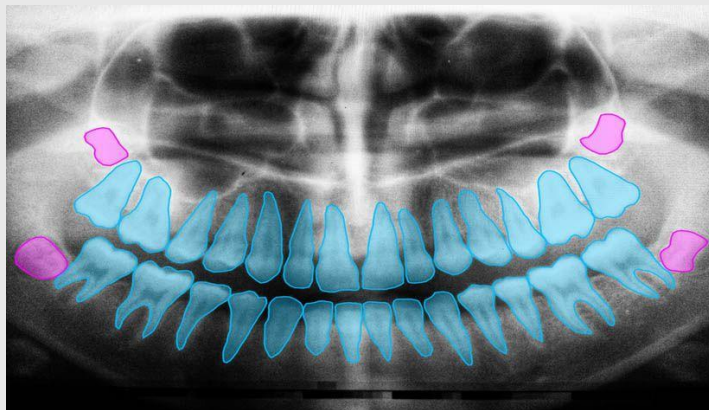
۱. تشخیص بیماری

۲. خودروهای خودران

۳. تغییر تصویر

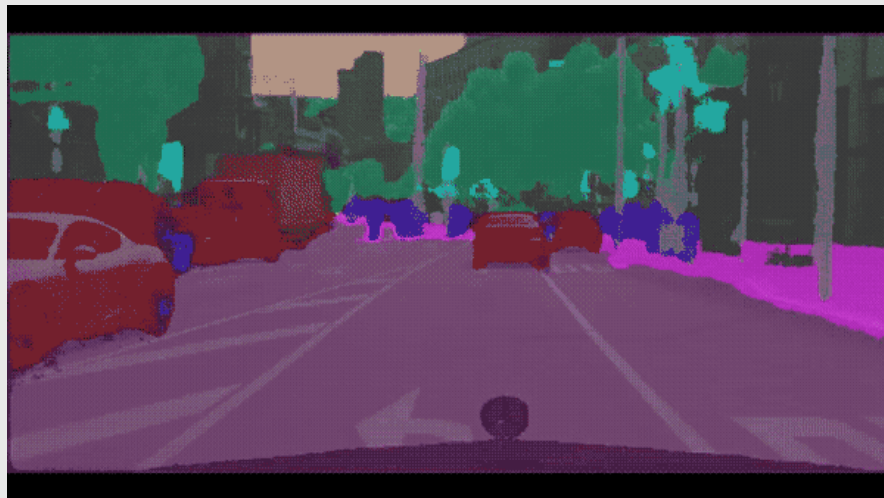
۴. تصویرهای هوایی

# تشخیص بیماری



- عکس OPG دندان و MRI
- تشخیص دندان‌های ناسالم

# خودروهای خودران



- شناسایی اشیاء مختلف
- شناسایی جاده

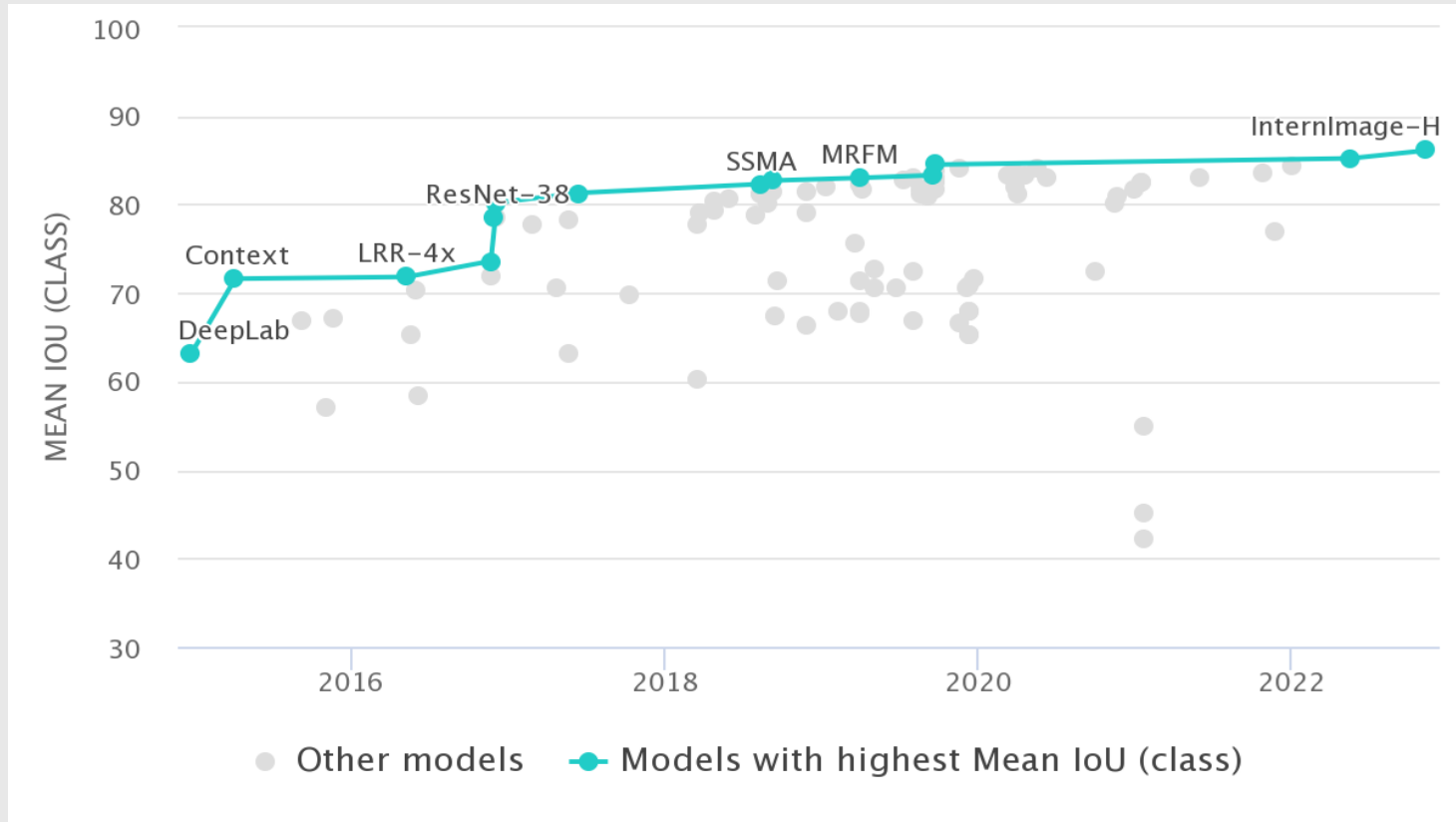
# مجموعه داده مناظر شهری



شکل ۲. نمونه‌ای از داده در شهر کلن

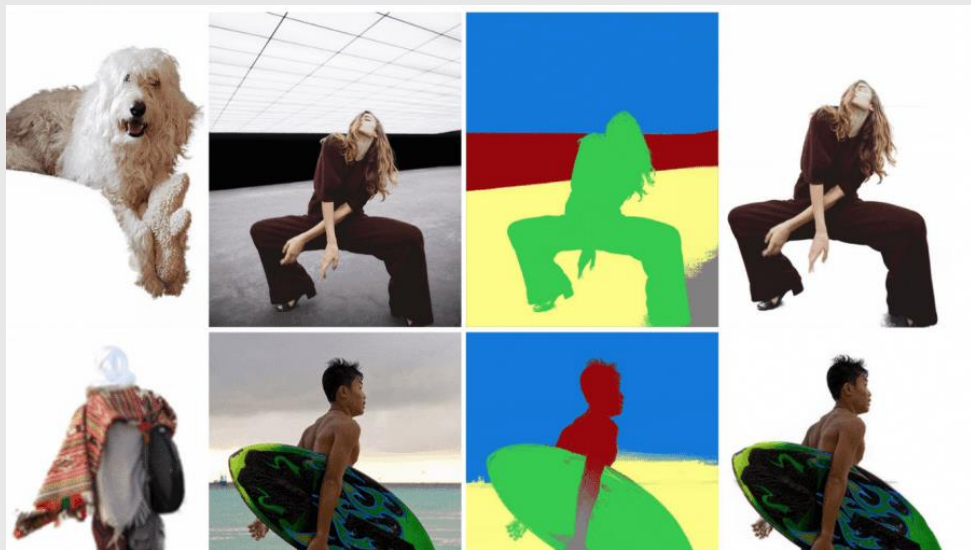


شکل ۱. نمونه‌ای از داده در شهر لینداو



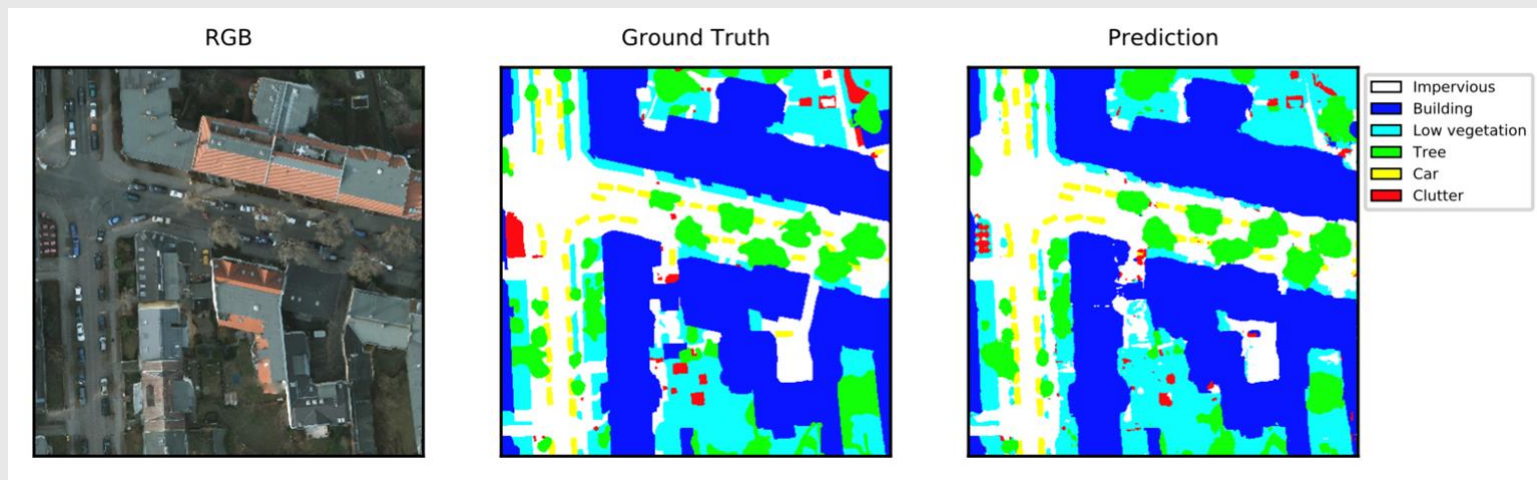


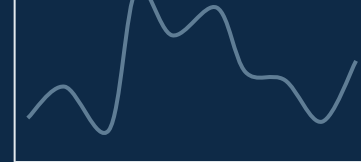
# تغییر تصویر



- تغییر رنگ یک دسته
- تغییر پس زمینه تصویر

# تصویرهای هوایی





# پرسش و پاسخ دیداری

## Visual Question Answering



# پرسش و پاسخ دیداری

- ترکیب بینش ماشین و پردازش زبان طبیعی
- پاسخ دادن به سوالات مربوط به یک تصویر

# پرسش و پاسخ دیداری

Who is wearing glasses?

man



woman



Where is the child sitting?

fridge



arms



Is the umbrella upside down?

yes



no



How many children are in the bed?

2



1



# کاربردهای پرسش و پاسخ دیداری

۱. تشخیص بیماری

۲. کمک به افراد نابینا

۳. تبلیغات

# تشخیص بیماری

## Selected Radiology Images



## Input Questions

Is there abnormality?

Submit

Yes

60%

No

20%

pneumonia

10%

atelectasis

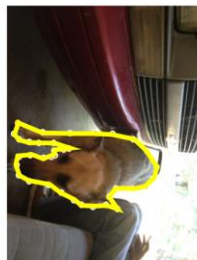
9%

cancer

1%



# کمک به افراد نابینا



Question: What is this?  
Answer: dog



Question: What does  
this package say?  
Answer: burrito



Question: Can you tell  
me what color this top  
is please?  
Answer: purple



Question: On this  
rectangular backup  
battery, how many  
lights are on?  
Answer: 2



Question: What is this?  
Answer: crystal



Question: What brand  
is this?  
Answer: chesters



Question: What color  
on left?  
Answer: grey



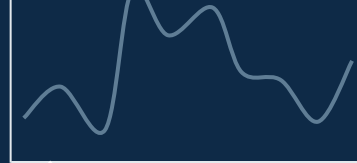
Question: How many  
tablets are in this box?  
Answer: 8

- عکس گرفتن از روبه‌رو
- سوال پرسیدن با صدا
- جواب گرفتن با صدا

# تبلیغات

سوال	جواب
با دیدن این تبلیغ باید چه کاری کنیم؟ چرا باید بر اساس این تبلیغ، این کار را بکنیم؟	باید کفش‌های Nike بخری. چون به شما هویت ورزشکاری می‌دهد.
با دیدن این تبلیغ باید چه کاری کنیم؟ چرا باید بر اساس این تبلیغ، این کار را بکنیم؟	باید این بازی کامپیوتری را بخری. چون با آن یک فوتبال واقعی را تجربه می‌کنید.
با دیدن این تبلیغ باید چه کاری کنیم؟ چرا باید بر اساس این تبلیغ، این کار را بکنیم؟	باید مراقب خشونت خانگی باشی. چون می‌تواند در روشنایی پنهان شود.





# شناسایی و ردیابی اشیا

## Object Detection



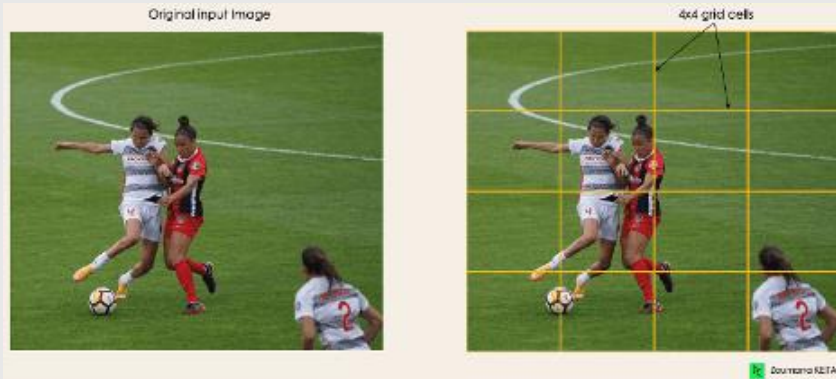
# شناسایی و ردیابی اشیا

• خروجی الگوریتم شناسایی اشیا:

- مختصات شی
- چرخش
- مقیاس



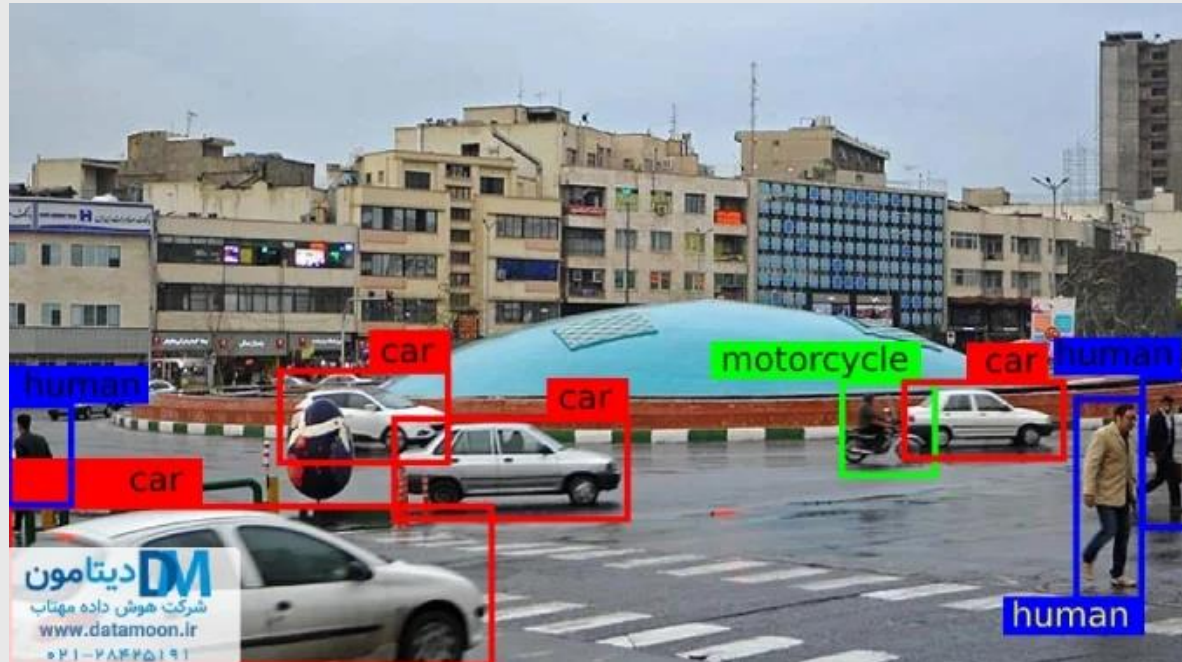
# Yolo



- الگوریتم پیشرفته تشخیص شی به شکل بلادرنگ
- همان You Only Look Once
- برتری‌ها نسبت به دیگر الگوریتم‌ها:
  - سرعت (Speed)
  - دقت تشخیص (Detection Accuracy)
  - تعمیم خوب (Good Generalization)
  - متن‌باز بودن (Open-source)



# شناسایی اشیای داخل خیابان

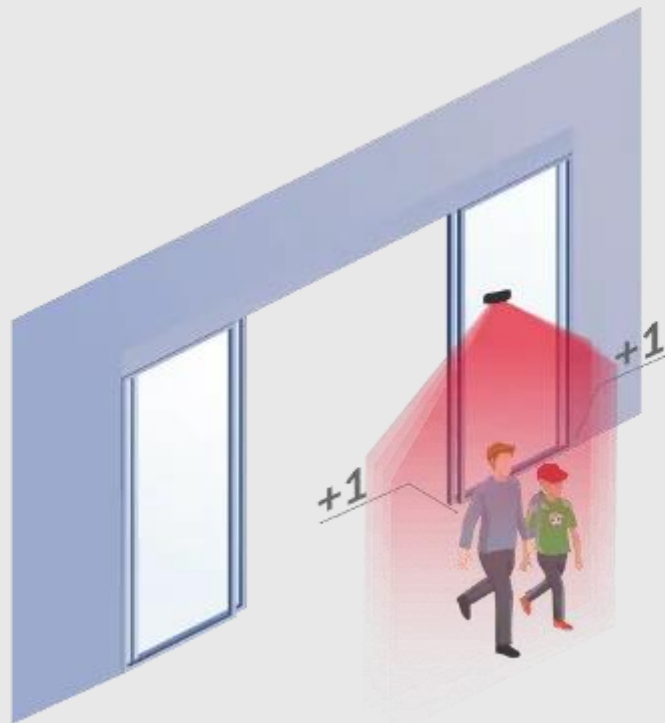


# تشخیص چهره



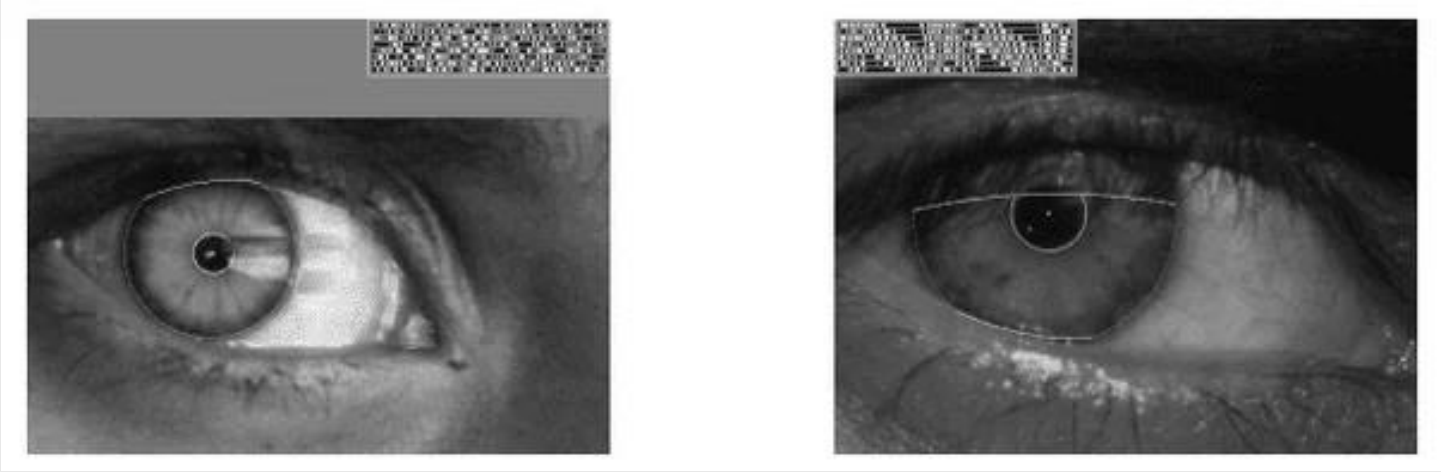
نام: شهره کسایی  
جنسیت: خانم  
مجوز دارد

# شمارش افراد





# تشخیص هویت از روی عنابه‌ی چشم



# استخراج شی از تصویر یا ویدیو



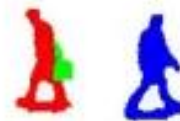
# تشخیص خنده



# رباتیک

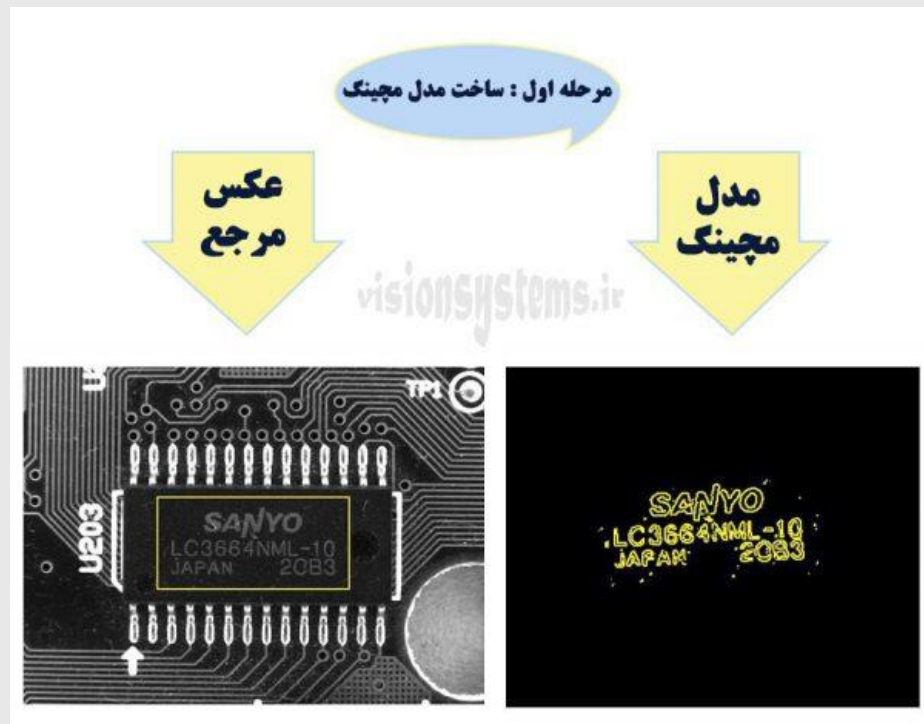


# ایده اصلی شناسایی اشیا در پردازش تصویر



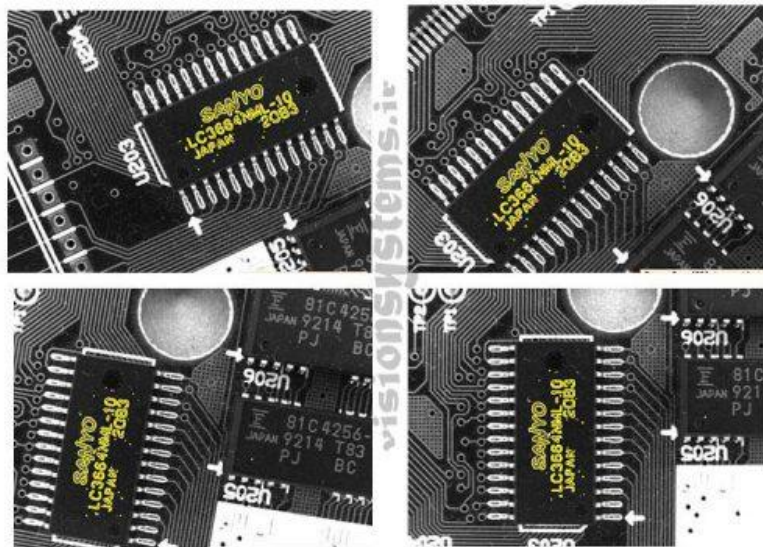


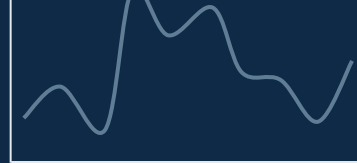
# ساخت مدل برای تشخیص اشیا



# شناسایی اشیا

مرحله دوم : پیدا کردن مدل





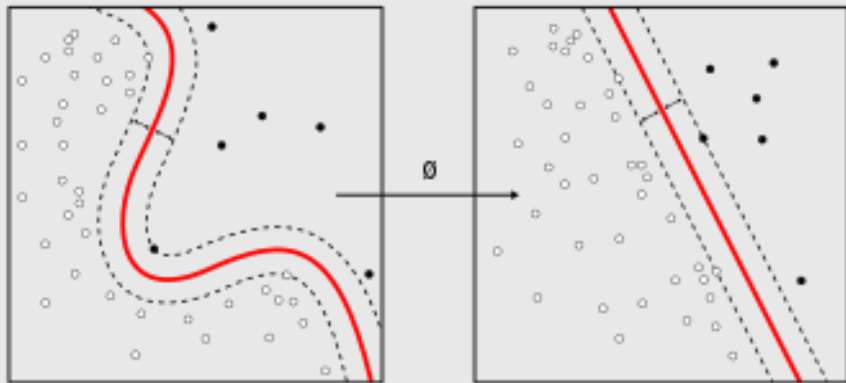
# مدل‌های تولیدکننده برای ساخت تصاویر

## Generative Models for Image Synthesis



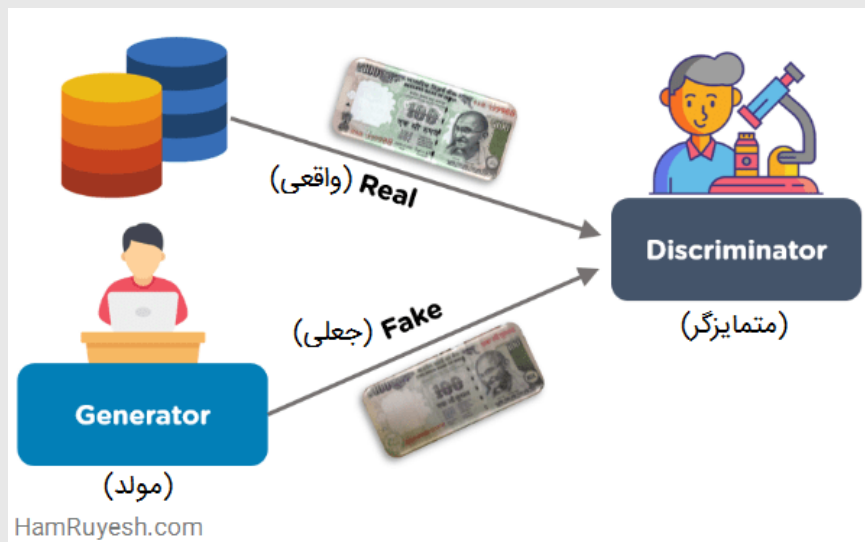


# مدل‌های انتشاری



- طراحی شده در سال ۲۰۱۵
- خراب کردن داده‌ها با اضافه کردن تدریجی نویز

# شبکه مولد تخاصمی (Generative Adversarial Network)



- طراحی شده در سال ۲۰۱۴
- شامل دو شبکه‌ی عصبی

# DALL-E ۲



- محصول آزمایشگاه تحقیقاتی OpenAI
- پیشروترین تولیدکننده متن به تصویر
- اولین نرم افزار با قابلیت کنترل فوق العاده بر ویژگی های تصاویر











# Stable Diffusion



- محصول StabilityAI و با همکاری EleutherAI و LAION
- ویژگی خاص شفافیت
- شامل کد منبع باز
- قادر به تولید تصاویر فوق العاده خلاقانه از افراد مشهور









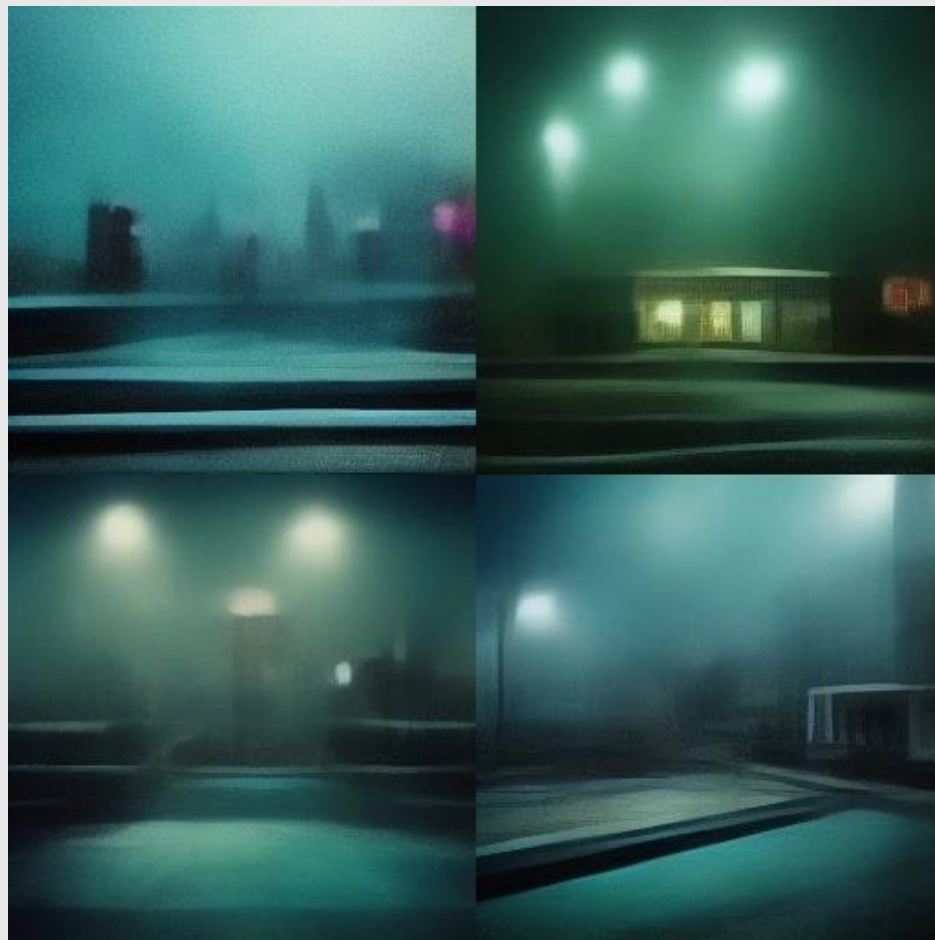




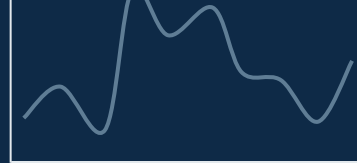
# میدجرنی – Midjourney



- ارائه از طریق یک سرور Discord
- غیرممکن برای ایجاد تصاویر «واقعی»
- رایگان برای ۲۵ تصویر اول، ۱۰ دلار در ماه برای ۲۰۰ تصویر بعدی
- قوانین میدجرنی:
  - استفاده نادرست از تصاویر نکنید.
  - از ربات برای تولید محتوای نامناسب استفاده نکنید.
  - به همه احترام بگذارید.



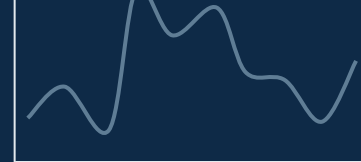




# جمع‌بندی





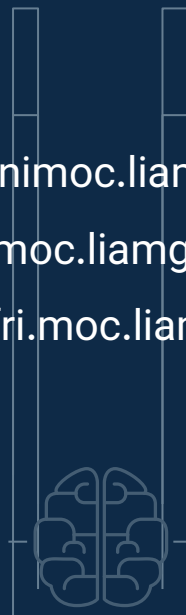


# ممنون از توجه شما

ارتباط با ما



absoltanimoc.liamg@۰۲  
imanmmoc.liamg@۱۳۸۱  
mahdi.jfri.moc.liamg@۷۹



## منابع

[https://en.wikipedia.org/wiki/History\\_of\\_the\\_camera](https://en.wikipedia.org/wiki/History_of_the_camera)

[https://en.wikipedia.org/wiki/Facial\\_recognition\\_system](https://en.wikipedia.org/wiki/Facial_recognition_system)

[https://en.wikipedia.org/wiki/Optical\\_character\\_recognition](https://en.wikipedia.org/wiki/Optical_character_recognition)

[https://en.wikipedia.org/wiki/Activity\\_recognition#Vision-based\\_activity\\_recognition](https://en.wikipedia.org/wiki/Activity_recognition#Vision-based_activity_recognition)

[https://en.wikipedia.org/wiki/Image\\_retrieval](https://en.wikipedia.org/wiki/Image_retrieval)

[https://en.wikipedia.org/wiki/Content\\_based\\_image\\_retrieval](https://en.wikipedia.org/wiki/Content_based_image_retrieval)

## منابع

[https://en.wikipedia.org/wiki/Video\\_tracking](https://en.wikipedia.org/wiki/Video_tracking)

[https://www.visionsystems.ir/matching\\_algorithm\\_image\\_processing/](https://www.visionsystems.ir/matching_algorithm_image_processing/)

[https://en.wikipedia.org/wiki/Object\\_detection](https://en.wikipedia.org/wiki/Object_detection)

<https://datamoon.ir/فناوری-تشخیص-و-شناسایی-اشیا/>

[https://en.wikipedia.org/wiki/Computer\\_vision](https://en.wikipedia.org/wiki/Computer_vision)

[https://en.wikipedia.org/wiki/Automatic\\_image\\_annotation](https://en.wikipedia.org/wiki/Automatic_image_annotation)

## منابع

<https://keymakr.com/blog/semantic-segmentation-uses-and-applications/>

<https://paperswithcode.com/sota/semantic-segmentation-on-cityscapes>

<https://paperswithcode.com/dataset/cityscapes>

<https://neurohive.io/en/state-of-the-art/image-editing-becomes-easy-with-semantically-meaningful-objects-generated/>

<https://visualqa.org/>

با تشکر از استاد محترم درس، سرکار خانم دکتر کسایی

ارائه‌ی مطالب علمی و فنی

دانشگاه صنعتی شریف

بهار ۱۴۰۲