



دانشگاه صنعتی امیرکبیر

(پلی تکنیک تهران)

دانشکده مهندسی برق

بینایی ماشین (Machine Vision)

Assignment 1

نگارش

علی دانشپور

محمد مهدی نجفی زاده

استاد

دکتر زهرا سادات شریعتمداری مرتضوی

مهر ۱۴۰۳

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

تشکیل تصویر و طراحی دوربین

۱ - در دوربین pinhole دریچه‌ای در نظر می‌گیرند تا تمام اشعه‌های بازتاب شده از آبجکت به صفحه سنسور برخورد نکند. کم و زیاد شدن اندازه این دریچه چه تاثیری در شکل‌گیری تصویر نهایی دارد؟

در دوربین pinhole، دریچه‌ای کوچک به نام اپرچر^۱ وجود دارد که نقش کلیدی در شکل‌گیری تصویر ایفا می‌کند. تغییر اندازه این دریچه تاثیر بسزایی بر روی خصوصیات تصویر نهایی دارد که در ادامه به برخی از آن‌ها اشاره می‌کنیم.

• وضوح تصویر

دریچه کوچکتر: تصویر واضح‌تر و شارپ‌تر می‌شود، زیرا پرتوهای نور کمتری وارد می‌شوند و تداخل کمتری ایجاد می‌کنند.

دریچه بزرگتر: تصویر تار و محو می‌شود، چون پرتوهای نور بیشتری از زوایای مختلف وارد می‌شوند.

• روشنایی تصویر

دریچه کوچکتر: تصویر تاریک‌تر می‌شود، زیرا نور کمتری به سنسور می‌رسد.

دریچه بزرگتر: تصویر روشن‌تر می‌شود، چون نور بیشتری وارد می‌شود.

• عمق میدان

دریچه کوچکتر: عمق میدان افزایش می‌یابد، یعنی محدوده بیشتری از صحنه در فوکوس خواهد بود.

دریچه بزرگتر: عمق میدان کاهش می‌یابد. به این معنا که فقط یک ناحیه مشخص (به عنوان مثال،

پس‌زمینه یا پیش‌زمینه) در حالت وضوح خواهد بود و سایر بخش‌ها ممکن است محو شوند. این ویژگی در عکاسی عموماً به منظور تأکید بر سوژه اصلی و جدا کردن آن از پس‌زمینه استفاده می‌شود.

• پدیده پراش نور

دریچه خیلی کوچک: احتمال وقوع پراش نور^۲ افزایش می‌یابد که می‌تواند باعث کاهش وضوح

تصویر شود.

¹Aperture

²Diffraction

• زمان نوردهی

دریچه کوچکتر: نیاز به زمان نوردهی طولانی‌تر دارد.
دریچه بزرگتر: امکان استفاده از زمان نوردهی کوتاه‌تر را فراهم می‌کند.
بنابراین بین اندازه دریچه و ویژگی‌هایی از تصویر که اشاره کردیم یک مصالحه¹ وجود دارد. انتخاب دقیق و بهینه اندازه این دریچه از اهمیت ویژه‌ای برخوردار است.

۲ - الف) عملکرد سنسورهای تصویر موجود در دوربین را بیان نمایید.

سنسورهای تصویر در یک دوربین اجزای مهمی هستند که وظیفه جذب نور و تبدیل آن به سیگنال‌های الکتریکی را برعهده دارند. سپس این سیگنال‌ها به تصاویر دیجیتالی تبدیل می‌شوند که روزانه مشاهده می‌کنیم. این سنسورها با شناسایی فوتون‌ها (ذرات نور) که از طریق لنز وارد دوربین می‌شوند، عمل می‌کنند. در اینجا به تفکیک برخی از عملکرد آن‌ها را معرفی می‌کنیم.

• تبدیل نور به سیگنال الکتریکی

سنسورهای تصویر نور وارد شده را به سیگنال الکتریکی تبدیل می‌کنند. هنگامی که نور بر روی سطح سنسور فرود می‌آید، پیکسل‌های سنسور آن را دریافت می‌کنند و به سیگنال الکتریکی تبدیل می‌کنند.

• تشکیل تصویر دیجیتال

بر اساس اطلاعاتی که توسط سنسورها به شکل سیگنال الکتریکی به دست آمده است، پردازشگر در دوربین این اطلاعات را به تصویر دیجیتال تبدیل می‌کند. این اطلاعات به صورت فایل عکس یا ویدیو ذخیره می‌شود.

• ثبت اطلاعات نوری

سنسورها اطلاعات نوری در ابعاد مختلف را به شکل پیکسل‌ها در کادرهای تصویری ثبت می‌کنند. هر پیکسل سنسور وظیفه ثبت نور مربوط به آن بخش از تصویر را دارد.

• مدیریت نویز

سنسورهای پیشرفته قابلیت کاهش نویز در شرایط نوری مختلف را دارند.

• فیلتر رنگ

برای گرفتن اطلاعات رنگی، اکثر سنسورهای تصویر از فیلترهای رنگی مانند فیلتر بایر استفاده می‌کنند که فیلترهای قرمز، سبز و آبی را روی پیکسل‌های جداگانه قرار می‌دهد. این مورد سنسورها را قادر می‌سازد تا با اندازه‌گیری شدت نور در هر یک از این رنگ‌ها، اطلاعات رنگ را بگیرند.

¹Trade-off

ب) دو نوع سنسور (CCD (Charge-coupled device و

CMOS (Complementary Metal-Oxide-Semiconductor) را با هم مقایسه و

مزایای هر کدام را بیان نماید.

سنسورهای CCD و CMOS دو نوع اصلی از سنسورهای تصویربرداری دیجیتال هستند که در دستگاه‌های مختلف از جمله دوربین‌های دیجیتال، میکروسکوپ‌ها و تلفن‌های همراه مورد استفاده قرار می‌گیرند. هر کدام از این دو نوع سنسور مزایا و معایب خاص خود را دارند که بر عملکرد و کارایی آن‌ها تاثیر می‌گذارد. در ادامه به مقایسه این دو نوع سنسور و بررسی مزایای هر یک می‌پردازیم.

سنسور CCD از مجموعه‌ای از فتو دیودها تشکیل شده است که نور را به بار الکتریکی تبدیل می‌کنند. بار الکتریکی تولید شده در هر پیکسل در یک خازن ذخیره می‌شود. در این سنسورها، بارهای الکتریکی ایجاد شده در هر پیکسل به صورت ترتیبی به یکدیگر منتقل می‌شوند تا در نهایت به یک تقویت‌کننده مشترک برسند که سیگنال را به تصویر دیجیتال تبدیل می‌کند. این ساختار باعث می‌شود که سنسور CCD به دلیل داشتن تقویت‌کننده‌های خارجی، تصاویر بسیار با کیفیت و با نویز پایین تولید کند.

سنسورهای CMOS بر پایه تکنولوژی ترانزیستورهای نیمه‌هادی کار می‌کنند. هر پیکسل دارای تقویت‌کننده و مبدل آنالوگ به دیجیتال (ADC) اختصاصی است که باعث می‌شود سیگنال مستقیماً به دیجیتال تبدیل شود. در سنسور CMOS، پیکسل‌ها به صورت مستقل عمل می‌کنند و به همین دلیل خوانش اطلاعات از آن‌ها سریع‌تر و با مصرف انرژی کمتر انجام می‌شود.

مزایای CCD :

کیفیت تصویر بالا: به دلیل انتقال دقیق بار الکتریکی، کیفیت تصویر بالاتری نسبت به سنسورهای CMOS دارند.

دقت رنگ خوب: عملکرد بسیار خوبی در بازتولید رنگ‌ها دارند.

حساسیت نوری بالا: به دلیل توانایی بالای جمع‌آوری نور، در محیط‌های با نور کم عملکرد بهتری دارند.

مصرف انرژی بیشتر: به دلیل نیاز به منابع انرژی برای انتقال بار در طول سنسور، مصرف انرژی بالاتری دارند.

سرعت پایین‌تر: سنسورهای CCD به دلیل سرعت پایین‌تر پردازش، در دوربین‌های ویدئویی با نرخ فریم بالا کمتر استفاده می‌شوند.

یکنواختی نور: CCD به دلیل داشتن تقویت‌کننده‌های خارجی، به صورت یکنواخت عمل می‌کند و تصاویر بدون اشکالات الکتریکی داخلی گرفته می‌شوند.

سنسورهای چندگانه: در برخی از دستگاه‌های حرفه‌ای، می‌توان از چندین سنسور CCD استفاده کرد تا از طول موج‌های مختلف نور به طور همزمان عکس بگیرند که این امکان با سنسورهای CMOS کمتر امکان‌پذیر است.

مزایای CMOS :

سنسورهای CMOS در سال‌های اخیر محبوبیت زیادی پیدا کرده‌اند و به دلیل تکنولوژی ساخت ساده‌تر و مقرون به صرفه‌تر، به طور گسترده‌ای در انواع دوربین‌ها استفاده می‌شوند. ویژگی‌های اصلی سنسورهای CMOS عبارتند از:

مصرف انرژی پایین‌تر: این سنسورها به دلیل معماری خاص خود، به انرژی کمتری نیاز دارند و این موضوع باعث افزایش عمر باتری دوربین‌ها می‌شود.

سرعت بالاتر: به دلیل اینکه هر پیکسل به طور مستقل خوانده می‌شود، سرعت پردازش بالاتری دارند که این امر در عکاسی متوالی و فیلم‌برداری با فریم بالا اهمیت دارد.

حساسیت کمتر به نور: در شرایط نوری کم، سنسورهای CMOS معمولاً عملکرد ضعیف‌تری نسبت به CCD دارند.

وضوح و کیفیت کمتر: به طور کلی، در سنسورهای CMOS به دلیل نویز بیشتر، ممکن است کیفیت تصویر پایین‌تری داشته باشند. البته فناوری‌های جدیدی برای کاهش این نقص‌ها در CMOS ارائه شده‌اند.

قیمت کمتر: سنسورهای CMOS به دلیل تکنولوژی ساخت ساده‌تر و امکان تولید انبوه با هزینه کمتر، ارزان‌تر از سنسورهای CCD هستند. این مزیت باعث شده تا در بسیاری از دوربین‌های دیجیتال و تلفن‌های هوشمند از CMOS استفاده شود.

انعطاف‌پذیری بیشتر: با توجه به ساختار سنسورهای CMOS، امکان پیاده‌سازی تکنولوژی‌های اضافی بر روی این سنسورها بیشتر است. مانند امکان پردازش تصاویر در خود سنسور، کنترل نواحی حساس به نور و همچنین کاهش اثرات ناشی از نور پس‌زمینه.

کاربردهای مناسب برای هر سنسور CCD: بیشتر در برنامه‌های علمی، پزشکی، تصویربرداری فضایی و سیستم‌های امنیتی که نیاز به کیفیت بالای تصویر دارند استفاده می‌شود.

کاربردهای مناسب برای هر سنسور CMOS: مناسب برای کاربردهای مصرفی عمومی مانند دوربین‌های دیجیتال خانگی، تلفن‌های همراه و کاربردهایی که نیاز به سرعت بالاتر و مصرف انرژی کمتر دارند.

نتیجه‌گیری: در نهایت، انتخاب بین سنسورهای CCD و CMOS به نیازهای خاص دستگاه و کاربردی که مورد نظر است بستگی دارد. اگر کیفیت تصویر و حساسیت نوری اولویت داشته باشد، سنسور CCD انتخاب مناسبی خواهد بود. اما اگر سرعت بالا، مصرف انرژی پایین و هزینه مناسب‌تر مورد نظر باشند، سنسور CMOS برتری خواهد داشت.

ج) با توجه به اینکه سنسورها تنها به نور حساس هستند و حساسیتی نسبت به رنگ ندارند، برای ثبت تصاویر رنگی چه مکانیسمی باید در نظر گرفته شود؟

برای ثبت تصاویر رنگی با استفاده از سنسورهای حساس به نور، مکانیسم‌های مختلفی به کار گرفته می‌شود که رایج‌ترین آن‌ها استفاده از فیلترهای رنگی است. در ادامه برخی از این فیلترها را معرفی می‌کنیم.

• فیلتر رنگی بایر^۱ :

ماتریسی از فیلترهای رنگی (معمولاً قرمز، سبز و آبی) است که روی سنسور دوربین قرار می‌گیرد. این فیلتر به صورت الگوی موزاییکی است که اجازه می‌دهد هر پیکسل تنها یکی از رنگ‌های قرمز، سبز یا آبی را ثبت کند. نسبت فیلترها معمولاً به صورت ۲ سبز، ۱ قرمز و ۱ آبی است، زیرا چشم انسان نسبت به نور سبز حساسیت بیشتری دارد. سپس، با استفاده از الگوریتمی به نام دیموزایکینگ^۲، داده‌های حاصل از سنسور پردازش می‌شوند تا تصویر رنگی نهایی تولید شود.

• فیلتر RGB :

در این سیستم، سه سنسور جداگانه برای ثبت سه رنگ اصلی RGB استفاده می‌شوند. هر سنسور به طور مستقل از دو سنسور دیگر، یکی از رنگ‌های اصلی را دریافت و ثبت می‌کند. تصویر رنگی نهایی به ترکیب داده‌های این سه سنسور بستگی دارد. این روش به دوربین‌ها این امکان را می‌دهد تا از یک سیستم سه سنسوره برای ثبت تصاویر استفاده کنند.

فیلترهای بایر ساده‌تر و از نظر اقتصادی به صرفه‌تر هستند در حالی که فیلترهای RGB در شرایطی که احتیاج به وضوح بالاتر در رنگ‌ها داریم، مناسب هستند.

۳ – سه فاکتور مهم در عکاسی، دیافراگم، شاتر و ISO می‌باشند که تحت عنوان exposure triangle شناخته می‌شوند. تأثیر این سه فاکتور در تهیه یک تصویر با وضوح خوب را توضیح دهید.

مثلث نوردهی یا exposure triangle از سه ضلع سرعت شاتر، دیافراگم و ایزو تشکیل شده است. هر کدام از اضلاع در میزان نوردهی دوربین و کیفیت تصویر مؤثر هستند. ضمن آن که این سه ضلع به یکدیگر وابسته می‌باشند و در صورت تغییر یکی از متغیرها، دو متغیر یا ضلع دیگر نیز باید تنظیم شوند. با

^۱Bayer Filter

^۲Demosaicing

تنظیم و تغییر هر سه متغیر، در نهایت نوردهی دوربین کنترل می شود. عمق میدان، میزان نویز و تاری یا شفاف بودن تصویر توسط این سه عنصر اصلی مشخص می شوند. به مثلث نوردهی، مثلث دوربین نیز می گویند.

• دیافراگم (Aperture)

دیافراگم اندازه دهانه‌ای است که نور از آن به سنسور دوربین وارد می شود. با تغییر قطر دریچه دیافراگم در مثلث نوردهی، میزان نور ورودی به سنسور دوربین نیز کم یا زیاد می شود. هرچه دریچه دیافراگم بزرگ‌تر باشد (عدد دیافراگم کوچک‌تر باشد) نور بیشتری وارد سنسور می شود؛ در نتیجه نور تصویر افزایش می یابد. از سوی دیگر، هر چه دریچه دیافراگم کوچک‌تر باشد (عدد دیافراگم بزرگ‌تر باشد) ورود نور به سنسور محدود می شود؛ بنابراین تصویر تاریک‌تر می شود.

همچنین دیافراگم بازتر عمق میدان کمتری دارد، که باعث می شود سوژه اصلی واضح‌تر و پس زمینه تاریک شود. از طرفی با کوچک کردن دریچه دیافراگم، عمق میدان عکس‌ها افزایش یافته و تمامی سوژه‌ها در تصویر واضح دیده می شوند. این مورد مناسب عکاسی از مناظر می باشد.

• سرعت شاتر (Shutter Speed)

سرعت شاتر، در واقع مدت زمانی است که تابش نور بر سنسور دوربین اتفاق می افتد. این فاکتور نیز تأثیرات مهمی بر روی عکس دارد.

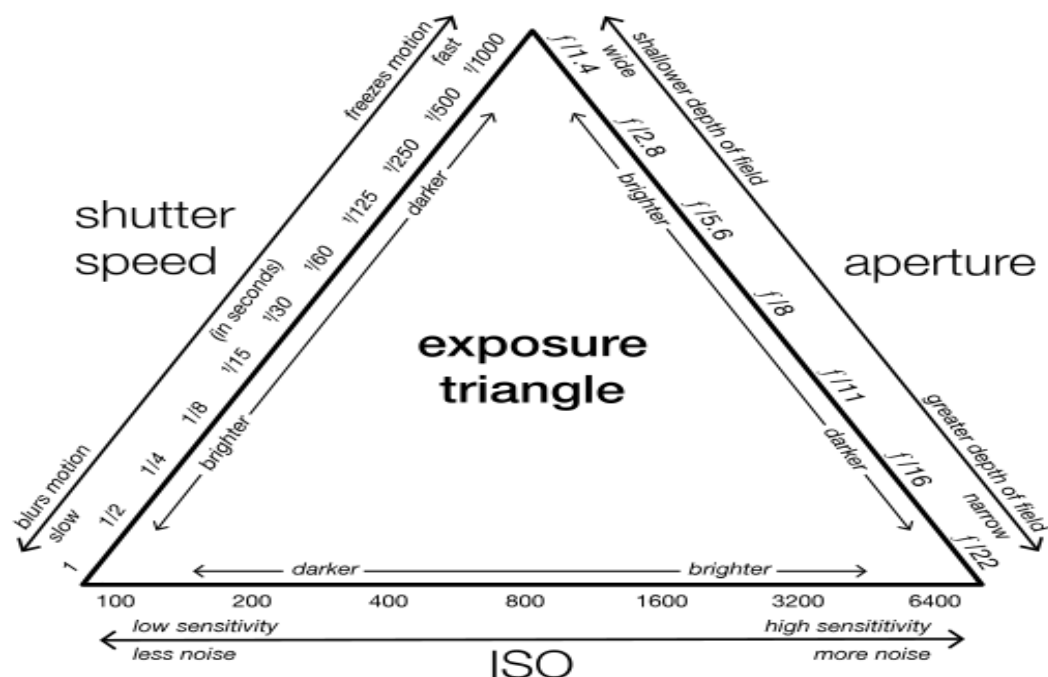
سرعت شاتر بالا حرکت را منجمد می کند و جزئیات دقیق را ثبت می کند. این مورد مناسب عکاسی از رویدادهای ورزشی است زیرا باعث انجماد حرکت می شود. از طرفی شاتر با سرعت پایین‌تر نور بیشتری وارد می کند و مناسب عکاسی در نور کم است، اما ممکن است باعث تاری حرکت شود. استفاده از سرعت شاتر پایین به عکاس این امکان را می دهد تا حرکت سوژه‌ها را ثبت کند. به عنوان مثال عکاسی از جریان آب با شاتر پایین، آب را به صورت صاف و نرم ثبت می کند. با این حال، برای چنین عکس‌هایی به دلیل اینکه دوربین در مدت زمان طولانی‌تری نور دریافت می کند، باید از سه پایه برای جلوگیری از لرزش دوربین استفاده شود.

• ایزو (ISO)

ایزو حساسیت سنسور دوربین به نور ورودی را نشان می دهد و به عکاس امکان می دهد تا تنظیمات نوردهی را بدون تغییر دیافراگم یا سرعت شاتر، کنترل کند. اعداد ایزو معمولاً از ۱۰۰ شروع می شوند و تا ۱۰۲۴۰۰۰ متغیر هستند. هر چه عدد انتخاب شده برای ایزو بزرگ‌تر باشد، میزان حساسیت سنسور دوربین افزایش یافته و هر چه عدد کوچک‌تر باشد، میزان حساسیت سنسور نیز کاهش می یابد.

ایزوی بالا، هرچند باعث افزایش روشنایی تصویر می شود، اما نویز بیشتری را نیز به تصویر اضافه می کند. به همین دلیل، برای داشتن تصاویر با وضوح بالا و نویز کمتر، استفاده از مقادیر ISO پایین مناسب‌تر است، مگر اینکه نور محیط ناکافی باشد و تنظیمات دیافراگم و شاتر محدود باشد.

در شکل ۱ مثلث نوردهی و وابستگی مقادیر شاتر، دیافراگم و ایزو به یکدیگر مشخص است.



شکل ۱: مثلث نوردهی (exposure triangle)

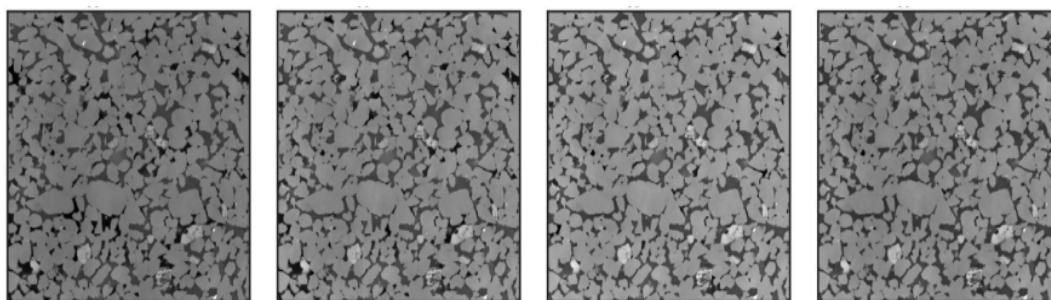
کاربرد بینایی ماشین

۱ - به دنیای اطراف خود نگاه کنید. ایده‌ای مطرح نمایید که با استفاده از بینایی ماشین قابل اجرا باشد. (منظور کاربردی است که هنوز مورد توجه دیگران قرار نگرفته است).

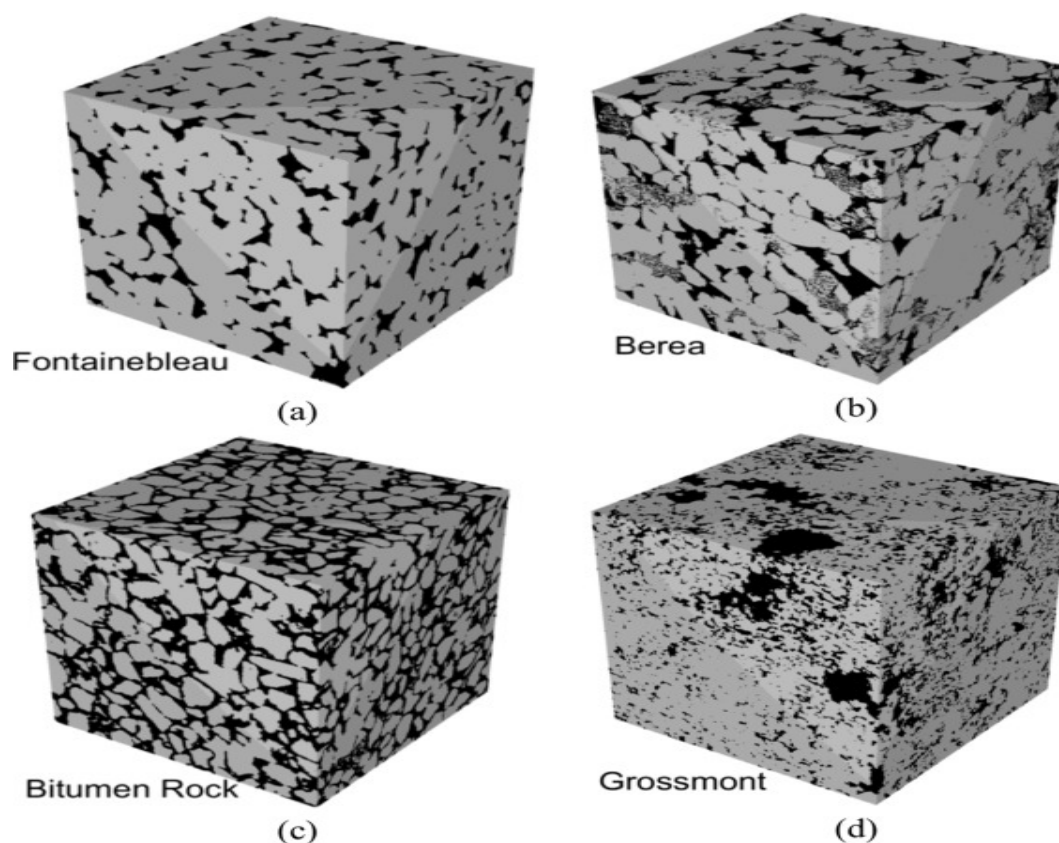
بینایی ماشین کاربرد گسترده‌ای در حوزه‌های تخصصی و زندگی روزمره دارد. با این حال با توجه به سوال مطرح شده در این تمرین، تنها به برخی کاربردهایی که کمتر مورد توجه قرار گرفته است اشاره می‌کنیم.

یکی از کاربردهایی که کمتر مورد توجه قرار گرفته است، پردازش تصاویر مستخرج از مخازن هیدروکربوری می‌باشد. با توجه به هزینه‌ی بالایی که فرآیند حفاری و استخراج نفت دارد، بنابراین شناخت دقیق مخازن نفتی از اهمیت ویژه‌ای برخوردار است. شناخت اولیه از این مخازن با توجه به تصاویری که از این مخازن به دست می‌آید، صورت می‌پذیرد. تا این لحظه بررسی و آنالیز این تصاویر در تمامی سکوه‌های نفتی به وسیله متخصصان زمین‌شناسی و مهندسی نفت صورت می‌گیرد. با استفاده از مدل‌های مبتنی بر یادگیری عمیق در حوزه پردازش تصاویر، امکان پردازش هوشمند این تصاویر و استخراج ویژگی‌های مورد نظر وجود دارد. با پردازش این تصاویر امکان پیش‌بینی خواص سنگ مخزن، شناسایی نقاطی که نفت در آن وجود دارد و ... وجود دارد. کاربردهای اشاره شده از اهمیت فوق العاده

زیادی در حوزه نفت، گاز و پتروشیمی دارند. مقالات [۹] و [۵] معدود مقالاتی می‌باشند که در این حوزه وجود دارد. نمونه ای از این تصاویر در اشکال ۲ و ۳ قابل مشاهده می‌باشد.



شکل ۲: نمونه از تصاویر مستخرج از مخازن هیدروکربوری



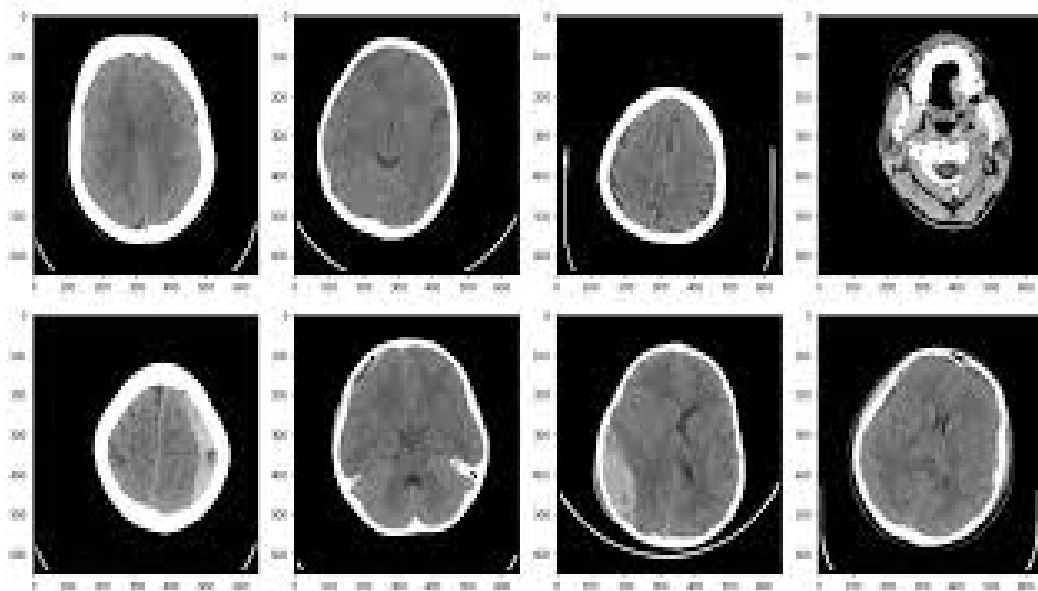
شکل ۳: نمونه از تصاویر مستخرج از مخازن هیدروکربوری

یکی دیگر از کاربردهایی که با وجود پژوهش‌هایی زیادی که در آن حوزه صورت گرفته است اما همچنان احتیاج به بهبودهای فراوانی دارد، تصاویر پزشکی می‌باشد. با توجه به ماهیت پیچیده این تصاویر و دقت فوق العاده بالایی که پردازش این تصاویر احتیاج دارد، می‌توانیم پژوهش‌های صورت گرفته تا این لحظه را تنها نقطه‌ی شروعی از مسیر پردازش تصاویر پزشکی در نظر گرفت.

با پردازش این تصاویر و استفاده از مدل‌های بینایی ماشین، امکان تشخیص بیماری‌ها، پیشگیری از وقوع بیماری‌ها، ارائه راهکارهای مناسب جهت درمان بیمار و ... وجود دارد. با توسعه و بهینه‌سازی مدل‌های مبتنی بر بینایی ماشین امکان بهبود روش‌های ارائه‌شده در این حوزه وجود دارد. مقالات [۱] و [۲] تنها تعداد محدودی از مقالات ارائه شده در این حوزه می‌باشند. نمونه‌ای از این تصاویر در اشکال ۴ و ۵ قابل مشاهده می‌باشد.



شکل ۴: نمونه‌ای از تصاویر پزشکی



شکل ۵: نمونه‌ای از تصاویر پزشکی

۲ - چگونه می‌توان از بینایی ماشین در جهت کمک به افراد با ناتوانی‌های جسمی کمک نمود؟

یکی از کاربردهای مهم هوش مصنوعی و به‌خصوص حوزه بینایی ماشین برطرف کردن نیازهای روزمره زندگی است. اما این موضوع برای افراد با ناتوانی‌های مختلف از اهمیت بیشتری برخوردار و البته چالش‌های بیشتری نیز دارد. سیستم‌های مبتنی بر بینایی ماشین کاربردهای گسترده‌ای در حل این مشکلات دارند که در ادامه به برخی از مهم‌ترین آن‌ها اشاره می‌کنیم.

یکی از چالش‌هایی که همواره برای افراد ناشنوا وجود دارد، ارتباط با مردم است. اکثریت جامعه تسلطی به زبان اشاره^۱ نداشته و یا خیلی محدود این زبان را می‌شناسند. همین موضوع کار ارتباط افراد ناشنوا با مردم را بیش از پیش دشوار می‌کند. استفاده از مدل‌های هوش مصنوعی و به ویژه مدل‌های حوزه بینایی ماشین امکان ترجمه زبان اشاره برای مردم را فراهم می‌کند. این مدل‌ها با پردازش بلادرنگ حرکات دست افراد ناشنوا، به ترجمه این حرکات و تبدیل آن‌ها به زبان طبیعی می‌پردازند. همچنین برخی از مدل‌ها با پردازش حرکات دست و استفاده از سیستم‌های توصیه‌گر، کلمات بعدی را به ناشنویان پیشنهاد می‌دهند که باعث راحتی و افزایش سرعت افراد ناشنوا می‌شود.

مقالات [۴، ۷، ۸]، تنها تعداد محدودی از مقالاتی می‌باشند که با توسعه مدل‌های حوزه بینایی ماشین به دنبال تسهیل ارتباط افراد ناشنوا با مردم می‌باشند.

همچنین سیستم‌های هشداردهنده مبتنی بر بینایی ماشین نیز یکی از ابزارهای بسیار مفید برای پیشگیری از حوادث و خطرات برای افراد نابینا یا کم‌بینا می‌باشد. برای افرادی که درک صحیحی از محیط اطراف خود ندارند، این سیستم‌ها می‌توانند به شناسایی موانع و خطرات کمک کرده و همچنین هشدارهای لازم را به این افراد بدهند.

با توجه به اینکه افراد نابینا و کم‌بینا توانایی خواندن متون را به راحتی ندارند، سیستم‌های تبدیل متن به گفتار^۲ می‌توانند کمک شایانی برای این افراد باشد. همچنین متاسفانه در معاملات مالی برخی از افراد به دنبال سودجویی از افراد نابینا هستند. مدل‌های مبتنی بر بینایی ماشین توانایی تشخیص تعداد سکه‌ها یا پول کاغذی و همچنین ارزش‌گذاری آن‌ها را دارند. این مورد می‌تواند جلوی سواستفاده از افراد نابینا و کم‌بینا را بگیرد.

¹Sign Language

²Text-to-Speech

۳ - مقاله‌ای به دلخواه انتخاب نمایید که در ارتباط با پردازش یکی از مشخصه‌های بیومتریک باشد. روش پیشنهادی مقاله انتخابی را حداکثر در دو صفحه خلاصه نمایید.

(الف) مقاله می‌تواند کنفرانس یا ژورنال باشد.

(ب) در مقاله از روش کلاسیک جهت پردازش استفاده کرده باشد (از روش‌های دیپ لرنینگ استفاده نکرده باشد).

(ج) سعی کنید مقاله‌ای انتخاب کنید که پس از سال ۲۰۱۶ چاپ شده باشد.

(د) برای پیدا کردن مقاله می‌توانید کلیدواژه خود را در سایت زیر جستجو نمایید.

www.connectedpapers.com

مقاله‌ای که برای این تمرین در نظر گرفتیم [۲] مربوط به سال ۲۰۱۹ می‌باشد. این مقاله در ژورنال International Journal of Electrical and Computer Engineering به چاپ رسیده است. عنوان مقاله Improving of fingerprint segmentation images based on K-means and DBSCAN clustering می‌باشد. به طور خلاصه این مقاله یک رویکرد جدید برای بخش‌بندی^۱ تصاویر اثر انگشت با استفاده از خوشه‌بندی K-means و روش DBSCAN معرفی می‌کند. بخش‌بندی تصویر اولین قدم در سیستم‌های تشخیص اثر انگشت است، زیرا اثر انگشت را از پس‌زمینه تصویر جدا می‌کند. بخش‌بندی دقیق عملکرد استخراج ویژگی و فرآیندهای تطبیق در مراحل بعدی را ساده‌تر می‌کند. در ادامه جزئیات روش ارائه شده در این پژوهش را به تفصیل شرح می‌دهیم.

به طور کلی رویکرد ارائه‌شده در این مقاله شامل سه بخش کلی می‌باشد. بخش اول پیش‌پردازش می‌باشد که از فیلترهای sobel و Top-Hat برای بهبود کیفیت تصاویر اثر انگشت استفاده می‌کند. فیلتر sobel برای تشخیص لبه‌ها و محاسبه گرادیان تصویر استفاده می‌شوند. فیلتر Top-Hat از عملیات مورفولوژیکی^۲ برای استخراج عناصر کوچک و جزئیات تصویر استفاده می‌کند.

بخش دوم مربوط به بخش‌بندی تصاویر می‌باشد. هر تصویر به بلوک‌هایی کوچک که با یک‌دیگر هیچ همپوشانی ندارند، تقسیم می‌شود. هر بلوک با یک بردار ویژگی پنج بعدی بازنمایی می‌شود. این بردار ویژگی شامل واریانس، اختلاف میانگین، انسجام گرادیان، ridge direction و energy spectrum می‌باشد.

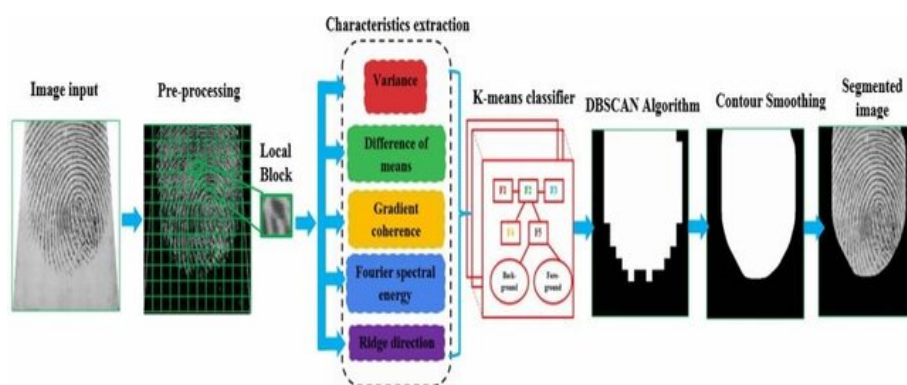
¹Segmentation

²Morphological

در این قسمت از روش خوشه‌بندی K-means به منظور تقسیم هر تصویر به پس‌زمینه^۱ و پیش‌زمینه (که همان اثر انگشت است) استفاده می‌شود.

در بخش‌هایی اقدام به برطرف کردن محدودیت‌های روش K-means می‌نماید. الگوریتم خوشه‌بندی DBSCAN برای بهبود طبقه‌بندی مورد استفاده قرار می‌گیرد. این الگوریتم به شناسایی و حذف نقاطی که نویزی بوده و اشتباه طبقه‌بندی شدند می‌پردازد. در انتها از Contour smoothing برای بهبود تصویر بخش‌بندی شده‌ی نهایی استفاده می‌شود.

این مقاله از مجموعه داده (FVC2004 (Fingerprint Verification Competition 2004) [۶] که شامل ۴ پایگاه داده (DB1، DB2، DB3، DB4) با انواع مختلف تصاویر اثر انگشت است، برای ارزیابی روش ارائه شده استفاده می‌کند. جزئیات رویکرد ارائه شده در مقاله در شکل ۶ مشخص است.



شکل ۶: نمای گرافیکی از ساختار روش پیشنهادی به منظور بخش‌بندی تصاویر اثر انگشت [۲]

به منظور ارزیابی مدل ارائه شده، این مقاله به مقایسه رویکرد ارائه شده با چهار تا روش‌های موجود پرداخته است. این چهار مدل به ترتیب SVM، K-Means با سه ویژگی، Morphological Processing و ACT (Automatic and Adaptive Thresholding) می‌باشند. رویکرد ارائه شده در مجموع عملکرد بهتری در مقایسه با تمام روش‌های اشاره شده از خود ارائه کرده است.

^۱Background

منابع و مراجع

- [1] Ait Skourt, Brahim, El Hassani, Abdelhamid, and Majda, Aicha. Lung ct image segmentation using deep neural networks. *Procedia Computer Science*, 127:109–113, 2018. PROCEEDINGS OF THE FIRST INTERNATIONAL CONFERENCE ON INTELLIGENT COMPUTING IN DATA SCIENCES, ICDS2017.
- [2] Cherrat, El Mehdi, Kh, Lily, and Bouzahir, Hassane. Improving of fingerprint segmentation images based on k-means and dbscan clustering. *International Journal of Electrical and Computer Engineering*, 9:2425–2432, 08 2019.
- [3] Gite, Shilpa, Mishra, Abhinav, and Kotecha, Ketan. Enhanced lung image segmentation using deep learning. *Neural Comput. Appl.*, 35(31):22839–22853, January 2022.
- [4] Juneja, Sapna, Juneja, Abhinav, Dhiman, Gaurav, Jain, Shashank, Dhankhar, Anu, and Kautish, Sandeep. Computer vision-enabled character recognition of hand gestures for patients with hearing and speaking disability. *Mobile Information Systems*, 2021:1–10, 12 2021.
- [5] Mahdaviara, Mehdi, Sharifi, Mohammad, and Rafiei, Yousef. Poreseg: An unsupervised and interactive-based framework for automatic segmentation of x-ray tomography of porous materials. *Advances in Water Resources*, 178:104495, 2023.
- [6] Maio, Dario, Maltoni, Davide, Cappelli, Raffaele, Wayman, Jim L., and Jain, Anil K. Fvc2004: Third fingerprint verification competition. In Zhang, David and Jain, Anil K.,

-
- editors, *Biometric Authentication*, pages 1–7, Berlin, Heidelberg, 2004. Springer Berlin Heidelberg.
- [7] Malik, M., Mansor, W., RASHID, NUR EMILEEN, and Rahman, M. Recognition of radar-based deaf sign language using convolution neural network. *International Journal of Integrated Engineering*, 15, 07 2023.
- [8] Obi, Yulius, Claudio, Kent, Budiman, Vetri, Achmad, Said, and Kurniawan, Aditya. Sign language recognition system for communicating to people with disabilities. *Procedia Computer Science*, 216:13–20, 01 2023.
- [9] Siavashi, Javad, Mahdaviara, Mehdi, Shojaei, Mohammad Javad, Sharifi, Mohammad, and Blunt, Martin J. Segmentation of two-phase flow x-ray tomography images to determine contact angle using deep autoencoders. *Energy*, 288:129698, 2024.