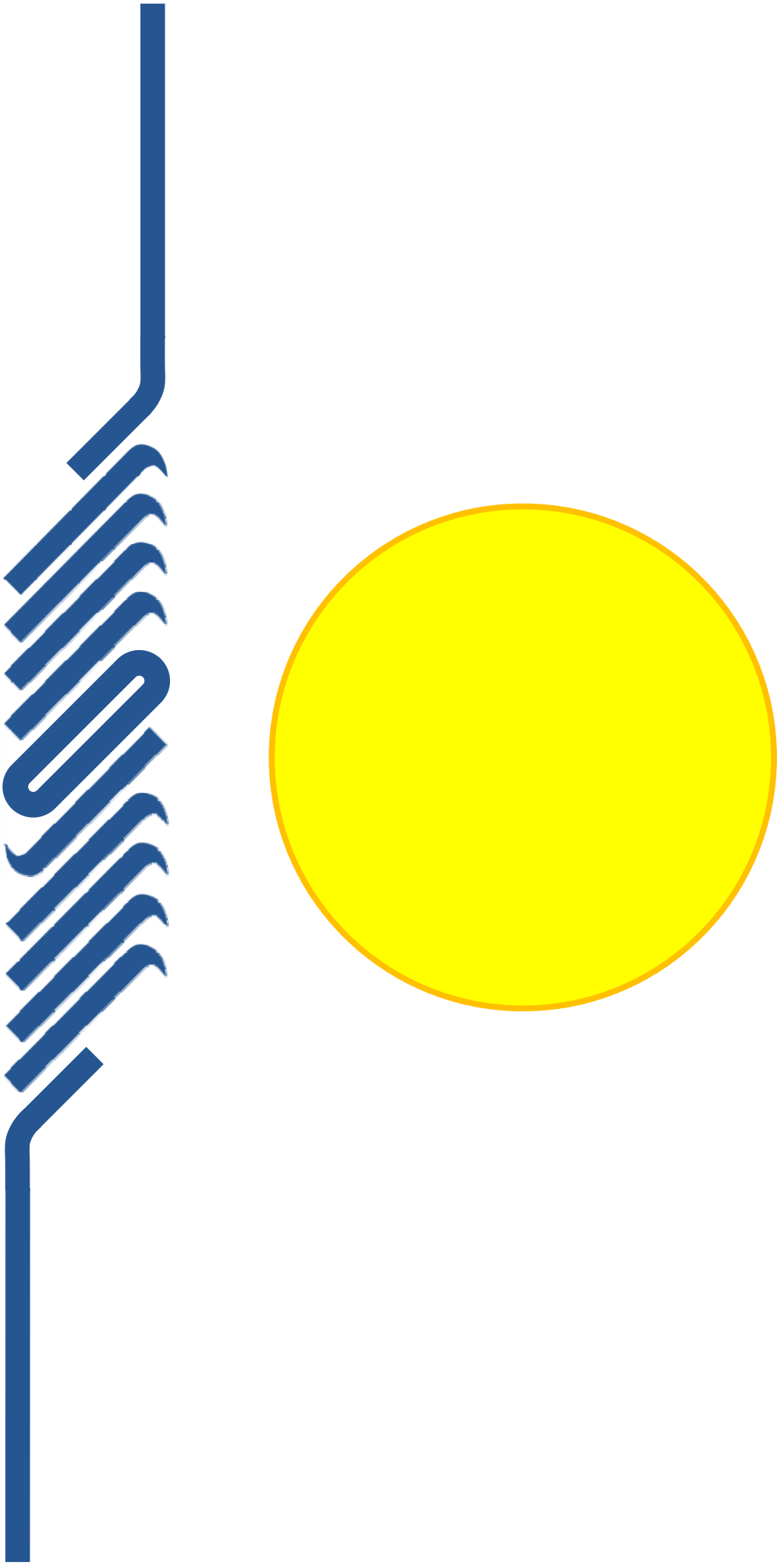


دانشکده فنی مهندسی

پایان نامه کارشناسی ارشد

رشته مهندسی پزشکی-بیوالکتریک

عنـــــوان:

بهبود مدل سلسله مراتبی بازشناسی اشیا مبتنی بر یافته‌های بیولوژیکی دستگاه بینایی

استاد راهنما : آقای دکتر محمد پویان

استاد مشاور: آقای دکتر رضا ابراهیم پور

نگارش : مسعود قـدرتی

زمستان ۱۳۹۰

کمترین تحریری از یک آرزو این است

آدمی را آب و نانی باید و آنگاه آوازی

در قناری‌ها نگه کن در قفس تا نیک دریابی

کزچه در آن تنگناشان باز، شادی‌های شیرین است

کمترین تحریری از یک زندگانی، آب، نان، آواز

ور فزونتر خواهی از آن، گاهگه پرواز

ور فزونتر خواهی از آن، شادی آغاز

ور فزونتر باز هم خواهی؟؟ بگویم باز...

آنچنان بر ما به نان و آب اینجا تنگسالی گشت

که کسی به فکر آوازی نباشد

اگر آوازی نباشد شوق پروازی نخواهد بود...

تقدیم به ایـــران، وطن م، پایدار و جاودان بمان.

تقدیر و تشکر

اساتید این خاک عاشقانی هستند که از حیث دانش همپای اساتید ممالک پیشرفته و از حیث تواضع بی‌همتا می‌باشند.کلمات، هر قدر پر مغز و پر معنا، در بیان قدردانی از آنان قاصراند. از تمامی این عزیزان بویژه جناب آقای دکتر محمد پویان و جناب آقای دکتر رضا ابراهیم‌پور، نهایت تشکر را دارم.

مسعود قدرتی

چکیده

انسان به‌سرعت و به‌طور موثر می‌تواند اشیا متفاوت را در صحنه‌های طبیعی و پیچیده شناسایی کند. این توانایی برجسته الهام‌بخش بسیاری از مدل‌های محاسباتی شناسایی اشیا بوده است. بسیاری از این مدل‌ها سعی به تقلید رفتار این دستگاه تحسین‌برانگیز دارند. دستگاه بینایی انسان اشیا را در یک سلسله مراتب از چندین مرحله پردازش شناسایی می‌کند. در طول این مراحل مجموعه‌ای از ویژگی‌ها با پیچیدگی فزاینده‌ای توسط بخش‌های مختلف دستگاه بینایی استخراج می‌گردد. ویژگی‌های ساده‌تر و ابتدایی مانند میله‌ها و لبه‌ها در سطوح اولیه مسیر بینایی پردازش شده و همچنان که در این مسیر بالا می‌رویم ویژگی‌های پیچیده‌تر مورد توجه و پردازش قرار می‌گیرند. این که چه ویژگی‌هایی از شی انتخاب می‌گردند و مورد پردازش قرار می‌گیرند همیشه مسئله‌ی مهمی در پردازش اشیا توسط قشر بینایی بوده است. برای پرداختن به این مساله، ما یک مدل سلسله مراتبی، که با یافته‌های بیولوژیکی سازگار است را گسترش داده و سپس آنرا در آزمایش‌های متفاوت شناسایی اشیا مورد ارزیابی قرار دادیم. در این مدل، مجموعه‌ای از بخش‌های شی، که با نام تکه‌های تصویر شناخته می‌شوند، در مراحل میانی استخراج می‌گردند. این تکه‌های تصویر برای آموزش در این مدل استفاده می‌شود و نقش مهمی در شناسایی شی ایفا می‌کنند. این تکه‌های تصویر بدون هدف و کورکورانه از مواضع مختلف از یک تصویر انتخاب شده که این موضوع می‌تواند منجر به استخراج تکه‌های غیر تبعیض‌آمیز گردد که در نهایت ممکن است عملکرد مدل را کاهش دهد. در روش پیشنهادی، از الگوریتم تکاملی برای انتخاب مجموعه‌ای از تکه‌های تصویر با اطلاعات مفید برای شناسایی اشیا گوناگون استفاده گردید. نتایج نشان می‌دهد که این تکه‌ها حاوی اطلاعات بیشتر و مفیدتری نسبت به تکه‌های معمول که به روش کورکورانه و تصادفی انتخاب می‌شوند هستند. همچنین نشان داده خواهد شد که قدرت مدل ارایه شده بر روی طیف وسیعی از تصاویر از پایگاه داده‌های مختلف قابل توجه بوده است. نتایج آزمایش‌ها نشان می‌دهد که ویژگی‌های انتخاب شده به‌طور کلی بخش‌های خاصی از تصاویر هدف هستند. این ویژگی‌های انتخاب شده که بخش‌هایی از اشیا هدف را ارایه می‌دهند مجموعه‌ای کارآمد برای تشخیص اشیا می‌باشند.

کلمات کلیدی: مدل بازشناسی اشیا، قشر بینایی، مدل سلسله مراتبی، الگوریتم‌های تکاملی، انتخاب ویژگی.

فهرست مطالب

عنوان **صفحه**

۱- مقدمه‌ای بر بازشناسی اشیا ۱

۱-۱- مقدمه ۱

۱-۲- مراحل شناسایی یک الگوی بصری ۲

۱-۳- مراحل محاسباتی درک تصویر ۴

۱-۳-۱- بینایی اولیه ۵

۱-۳-۲- بینایی لایه‌ی میانی ۵

۱-۳-۳- بینایی سطح بالا ۷

۱-۴- بازشناسی شی ۷

۱-۴-۱- بازشناسی اشیا در قشر مغز ۸

۱-۵- خلاصه فصل ۱۰

۲- دستگاه بینایی نسان ۱۲

۲-۱- مقدمه ۱۲

۲-۲- قرارگرفتن تصویر در شبکیه ۱۳

۲-۲-۱- سلول‌های دریافت‌کننده نور ۱۵

۲-۲-۲- سلول‌های گانگلیون ۱۷

۲-۲-۳- هسته‌های خمیدهی جانبی ۱۹

۲-۲-۴- سلول‌های ساده ۲۰

۲-۲-۵- سلول‌های پیچیده ۲۲

۲-۳- مسیر بینایی ۲۳

۲-۴- ساختار قشر بینایی مغز ۲۶

۲-۴-۱- ناحیه V1 ۲۹

۲-۵- ناحیه‌های بینایی در سطوح بالاتر ۳۳

۲-۶- خلاصه ۳۴

۳- مروری بر مدل‌ها ۳۶

۳-۱- مقدمه ۳۶
۳-۲- مروری بر تاریخچه مدل‌سازی سلول عصبی ۳۷
۳-۳- اولین مدل از سلسه مراتب سلول‌ها در قشر بینایی ۳۸
۳-۴- مدلی از چندین لایه از سلول‌های ساده و پیچیده(NEOCOGNITRON) ۳۹
۳-۴-۱- ساختار شبکه ۴۰
۳-۴-۲- خود سازماندهی شبکه ۴۴
۳-۵- مدل VISNET ۴۵
۳-۶- مدل‌سازی مغز در سطح ریزمداری (LAMINART) ۵۲
۳-۷- مدل HMAX ۵۷
۳-۷-۱- سلول‌ها یا واحدهای ساده و پیچیده ۶۱
۳-۷-۲- مرحله‌ی یادگیری ۶۵
۳-۷-۳- گام طبقه‌بندی ۶۶
۳-۷-۴- عملکرد مدل در مقایسه با انسان ۶۷
۳-۷-۴-۱- آزمایش اول ۶۷
۳-۷-۴-۲- آزمایش دوم ۷۲
۳-۷-۴-۳- آزمایش سوم ۷۳
۳-۸- خلاصه ۷۵

۴- مدل ارایه شده، نتایج و آزمایش‌ها ۷۷

۴-۱- مقدمه ۷۷
۴-۲- مواد و روش‌ها ۸۰
۴-۲-۱- شبیه‌سازی اولیه مدل ۸۰
۴-۲-۲- مقدمه الگوریتم ژنتیک ۸۳
۴-۳- روش پیشنهادی ۸۵
۴-۳-۱- ساختار کروموزوم ۸۵
۴-۳-۲- تابع شایستگی ۸۶
۴-۳-۳- معماری سیستم ۸۷
۴-۴- معرفی پایگاه داده ۸۸
۴-۵- طراحی آزمایش ۹۱
۴-۶- نتایج ۹۲

۴-۷- خلاصه ۱۰۲

۵- نتیجه‌گیری و پیشنهادات **۱۰۴**

۵-۱- نتیجه‌گیری ۱۰۴

۵-۲- پیشنهادات ۱۰۶

۶- مراجع **۱۰۷**

فهـرست شـکل‌ها

عنوان شکل **صفحـه**

فصل دوم

شکل(۲-۱). نمایش قسمت‌های مختلف چشم[۶۰]..... ۱۴

شکل(۲-۲). ترتیب قرارگیری سلول‌ها در شبکیه[۶۱]..... ۱۵

شکل(۲-۳). دو نوع گیرنده‌های نوری را نشان می‌دهد[۵۸]..... ۱۶

شکل(۲-۴). سلول‌های دوقطبی، سلول‌های افقی و سلول‌های آماکرین[۵۸]..... ۱۷

شکل(۲-۵) . سلول‌های گانگلیونی به تفاوت روشنایی در میان مرکز و محیط ناحیه گیرنده با کیفیت عالی پاسخ می‌دهند[۵۸]. ۱۹

شکل(۲-۶). پاسخ سلول ساده در مقابل تغییر زاویه محرک در میدان دریافت‌کنندگی[۵۸]. ۲۰

شکل(۲-۷). نمای شماتیکی از سلول‌های ساده و پیچیده و ترتیب قرارگیری آنها[۵۸]. ۲۱

شکل(۲-۸). تفاوت در پاسخ سلول‌های ساده و پیچیده، ناحیه خاکستری تیره میدان دریافت‌کنندگی سلول است[۵۸]..... ۲۲

شکل(۲-۹). میدان بینایی که از دو منطقه دید تشکیل شده است[۶۱]..... ۲۴

شکل(۲-۱۰). محل پردازش در قشر بینایی مغز با توجه به میدان‌های دید[۶۱]..... ۲۵

شکل(۲-۱۱). سمت راست، برش مقطعی مغز که قسمت تیره رنگ قشر مغز است. سمت چپ، قشر و لب‌های مختلف مغز[۵۹]. ۲۷

شکل(۲-۱۲). نواحی مختلف در قشر بینایی[۵۹]. ۲۸

شکل(۲-۱۳). دو مسیر در مغز وجود دارد. مسیر (۱) شی را تشخیص می‌دهد و مسیر (۲) مکان شی را تشخیص می‌دهد. ۲۸

شکل(۲-۱۴). V1 ناحیه ۱۷ قشر اولیه بینایی در تقسیم‌بندی برادمن[۵۹]. ۳۰

شکل(۲-۱۵). ناحیه دریافت‌کنندگی یک سلول ساده. خط تیره بالای هر پاسخ ثبت شده نشان دهنده طول مدت شدت روشنایی در هر حالت می‌باشد. سلول به خط نور عمودی که در مرکز ناحیه دریافت قرار دارد بیشترین پاسخ را می‌دهد. نقاط نوری پاسخ ضعیفی می‌دهند یا اصلاً پاسخی نمی‌دهند[۵۹]..... ۳۱

شکل(۲-۱۶). ناحیه دریافت‌کنندگی یک سلول ساده. خطوط منفی (-) نشان دهنده نقاط بازدارنده و خطوط مثبت (+) نشان دهنده نقاط وادارنده می‌باشد[۵۸]. ۳۲

شکل(۲-۱۷). ستون‌های جهت در سلول V1 از قشر بینایی یک میمون (با تاباندن نور). رنگ‌های مختلف نشان دهنده جهت‌های مختلف می‌باشند. همچنین رنگ قرمز و سبز نشان دهنده بیشترین پاسخ فعالیت نورون به نور در جهت افقی و عمودی می‌باشد[۶۰]..... ۳۳

فصل سوم

شکل(۳-۱). مدل ساده نرونی که توسط HODGKIN و HUXLEY ارایه گردید[۸]. ۳۸

شکل(۳-۲). مدل سلسله مراتبی HUBEL و WIESEL برای ایجاد سلول‌های ترکیبی از سلول‌های ساده[۵۸]. ۳۹
شکل(۳-۳). ساختار کلی سلسله مراتب در NEOCOGNITERON. [۱۰]. ۴۰
شکل(۳-۴). لایه‌های شبکه، لایه‌های ساده S (U_{SL}) و پیچیده C (U_{CL}) پشت سر هم قرار گرفته‌اند [۱۰]. ۴۱
شکل(۳-۵). ناحیه دریافت سلول‌ها و نحوه اتصالات [۱۰]. ۴۱
شکل(۳-۶). نحوه عملکرد مدل به تغییر مکان شی در تصویر[۱۰]. ۴۳
شکل(۳-۷). ساختار شبکه VISNET .[۱۴]..... ۴۶
شکل(۳-۸). نمایش یادگیری بر پایه تغییرات مداوم در برابر جابه‌جایی[۱۵]..... ۴۹
شکل(۳-۹). نمایش محرک به‌صورتی که دارای تغییرات جزئی در ارایه به شبکه هستند[۱۵]. ۵۲
شکل(۳-۱۰). نحوه ادراک مغز از مربع KANIZSA SQUARE، AوB) مغزما وجود مربعی را در شکل تشخیص می‌دهد که در واقع وجود ندارد. C) شبیه سازی نحوه درک مغز از شکل B [۱۹]. ۵۳
شکل(۳-۱۱). مدارهای نورنی از شبکیه تا V1 [۱۸]..... ۵۴
شکل(۳-۱۲). مدارهای کامل LGN، V1 و V2 [۲۰]. ۵۵
شکل(۳-۱۳). وجود شفافیت B)عدم وجود شفافیت C) مرزهای دو شکل A و B[۲۱]. ۵۶
شکل(۳-۱۴). مرز ناحیه BD در سیستم رقابتی از مرز ناحیه AC می‌برد، در نتیجه BD تقویت و AC تضعیف می‌شود [۲۱]..... ۵۷
شکل(۳-۱۵). ساختار مسیر پردازش بینایی که تقریبا مورد قبول اکثر محققان می‌باشد[۳۰]. ۵۸
شکل(۳-۱۶). قسمت‌های درگیر در قشر بینایی مغز حین شناسایی یک شی با توجه به پیچیدگی شی. ۵۹
شکل(۳-۱۷). ساختار کلی و لایه‌های مختلف مدل و بخش‌های مغزی موجود در قشر بینایی [۳۰]..... ۶۱
شکل(۳-۱۸). میدان دریافت کنندگی سلول‌های ساده S1 (حاصل از تابع گابور)[۵۰]..... ۶۲
شکل(۳-۱۹). ایجاد تغییرناپذیری نسبت به اندازه و موقعیت با استفاده از عملگر MAX در گذر از لایه S1 به C1 [۵۰]. ۶۳
شکل(۳-۲۰). ردیف بالا، تصاویر حیوان، ردیف پایین تصاویر غیر حیوان. ۶۸
شکل(۳-۲۱). نمایش از داده‌ها و نحوه نشان دادن تصاویر در آزمایش دسته‌بندی تصاویر حیوان از غیرحیوان بین مدل و بیننده انسانی[۳۰]. ۷۰
شکل(۳-۲۲). آزمایش ۱: مقایسه میان مدل و بیننده انسانی با حالت‌های مختلف ماسک[۳۰]. ۷۱
شکل(۳-۲۳). آزمایش ۲: جزئیات مقایسه میان مدل و بیننده انسانی. A) عملکرد مدل در برابر سطح عملکرد انسان. B,C,D,E) مقایسه میان بیننده انسانی و مدل برای هر تصویر. اعداد بالای تصاویر با رنگ سبز عملکرد مدل و رنگ سیاه عملکرد بیننده انسانی[۳۰]. ۷۳.
شکل(۳-۲۴). آزمایش ۳: مقایسه میان مدل و بیننده انسانی برای تصویر چرخیده شده[۳۰]. ۷۴

فصل چهارم

شکل(۴-۱). نتایج حاصل از پیاده‌سازی اولیه مدل. خروجی لایه اول (S1) ۶۴ تصویر فیلتر شده است (۱۶ اندازه در ۴جهت)، سپس این ۶۴ فیلتر دوبه‌دو بیشینه‌گیری می‌گردند و خروجی لایه بعد (C1) را می‌سازند. در این مرحله تکه‌های تصویر استخراج شده و ذخیره

می‌گردند. مرحله بعد محاسبه فاصله این تکه‌ها با تصاویر ورودی است که در نهایت خروجی لایه سوم (S2) را می‌سازد، خروجی این لایه بیشینه‌گیری شده و خروجی نهایی را تولید می‌کند. ۸۲
شکل(۴-۲). بیان تصویری الگوریتم ژنتیک، عملگر ادغام و جهش. ۸۴
شکل(۴-۳). ساختار کروموزوم و نحوه اشاره به ویژگی‌ها. سطر اول، ساختار دودویی از یک کروموزوم را نشان می‌دهد و ردیف دوم تکه‌هایی را که در مرحله یادگیری شرکت می‌کنند را نشان می‌دهد. ۸۶
شکل(۴-۴). نمایش بلوک دیاگرام مراحل مختلف در مدل پیشنهادی..... ۸۸
شکل(۴-۵). نمونه‌هایی از داده‌های استفاده شده از پایگاه دانشگاه صنعتی کالیفرنیا [۳۲]. ۸۹
شکل(۴-۶). نمونه‌هایی از پایگاه داده (B) GRAZ-01 و GRAZ-02 (A) [۵۳]. ۹۱
شکل(۴-۷). مقایسه عملکرد مدل برای تعداد مختلف داده آموزش. ۹۴
شکل(۴-۸). تصویر ماتریس جمعیت اولیه و جمعیت نهایی..... ۹۶
شکل(۴-۹). تکه‌های اولیه و انتخاب شده. نمونه‌های تصاویر ورودی (A)، ویژگی‌های اولیه (B)، ویژگی‌های بهینه (C). ۹۷
شکل(۴-۱۰). نرخ انتخاب تکه‌های تصویر با اندازه‌های مختلف در تصاویر متفاوت. ۹۹
شکل(۴-۱۱). مقایسه عملکرد مدل برای تعداد مختلف ویژگی‌های انتخاب شده در تصاویر مختلف. ۱۰۰
شکل(۴-۱۲). منحنی ROC برای تصاویر مختلف پایگاه داده GRAZ-02 و مقایسه نتایج مدل با دیگر روش‌ها. ۱۰۲

فهرست جـدول‌ها

عنوان جدول **صفحـه**

فصل سوم

جدول(۳-۱). بعد هر لایه و تعداد اتصال‌ها در شبکه VisNet.[۱۵]..... ۴۷

جدول(۳-۲). پارامترهای لایه‌های S1 و C1 [۲۸]. ۶۴

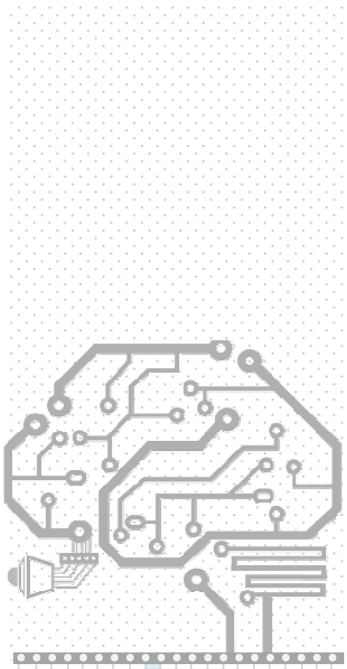
فصل سوم

جدول(۴-۱). تعداد و اندازه تصاویر پایگاه داده دانشگاه صنعتی کالیفرنیا[۳۲]. ۹۰

جدول(۴-۲). مقایسه نتایج به‌دست آمده با [۲۸] روی پایگاه داده دانشگاه صتعتی کالیفرنیا. ۹۳

جدول(۴-۳). نتایج به‌دست آمده برای مجموعه داده GARZ-01 نتایج متوسط ۲۰ اجرای مستقل تصادفی می‌باشد. ۱۰۱

جدول(۴-۴). نتایج به‌دست آمده برای مجموعه داده GARZ-02 نتایج متوسط ۲۰ اجرای مستقل تصادفی می‌باشد. ۱۰۱



فصل اول

مقدمه ای بر بازشناسی اشیا

۱-۱- مقدمه

بازشناسی اشیا[1] همواره از اهمیت بسیار بالایی برخوردار بوده است. تاکنون تحقیقات بسیار زیادی پیرامون این موضوع صورت گرفته است، مسئله بازشناسی اشیا موضوع پیچیده‌ای است که سال‌ها مورد توجه و علاقه مهندسان کامپیوتر، الکترونیک و نیز محققان علوم عصب‌شناسی[2] بوده است. به‌طور کلی، بازشناسی اشیا به‌معنای یافتن یک شی در یک صحنه یا تصویر است. انسان و بسیاری از پستانداران قادر هستند که بسیاری از اشیا پیرامون خود را بدون کوچک‌ترین مشکلی شناسایی و تفکیک نمایند، هرچند ممکن است که اشیا موجود در یک صحنه در حالت‌های گوناگون و با زوایای دید مختلف و همچنین در اندازه‌های متفاوت باشند. حتی انسان قادر است اشیا را در حالتی که بخش‌هایی از آنها را نمی‌بیند، یا شی دیگری در مسیر دیدش قرار گرفته را نیز شناسایی کند. اگرچه این امر به‌نظر برای انسان و پستانداران بسیار ساده و بدون تلاش و صرف زمان زیاد است، اما در نوع خود یک فرآیند محاسباتی بسیار مشکل و پیچیده است. این موضوع یعنی حل و شناخت چگونگی شناسایی اشیا در مغز کلیدی ارزشمند برای پاسخ به بسیاری از سوالات علوم شناختی[3] و عصبی می‌باشد که می‌تواند منجر به ساخت ماشین‌های هوشمند در آینده گردد.

در این فصل تلاش شده است که به تعریف مسئله شناسایی اشیا یا به بیان دیگر بازشناسی اشیا بپردازیم و مفاهیم ابتدایی در این حوزه را به‌طور خلاصه بیان کنیم.

---

[1] Object Recognition
[2] Neurologists
[3] Cognitive Science

۱-۲- مراحل شناسایی یک الگوی بصری

الگوهای بصری را می‌توان به‌صورت ترکیبی از مولفه‌های تصویری[۱] که هر کدام از آنها دارای سطح روشنایی[۲] خاص خود هستند، در نظر گرفت. هدف کلی از شناسایی خودکار الگوهای بصری، انتساب نمونه‌ای از یک الگو که سیستم قبلاً آن را تجربه نکرده است به‌یکی از الگوهایی که از پیش برای سیستم معرفی شده است، می‌باشد. این انتساب بر اساس تحلیل ویژگی‌های استخراج شده از الگوی ورودی و کلاس‌های موجود انجام می‌گیرد. رسیدن به این هدف کار مشکلی است، زیرا ممکن است الگوی جدید (تصویر ورودی) نسبت به نمونه‌های قبلی تغییرات زیادی داشته باشد. این تغییرات می‌توانند ناشی از شرایط محیطی در هنگام تهیه‌ی تصویر و یا مربوط به تغییرات اجتناب‌ناپذیر در خود الگو باشند. از جمله این تغییرات می‌توان به اختلال (نویز) در وسیله‌ی گیرنده‌ی تصویر و یا تار بودن تصویر در اثر تنظیم نبودن دوربین اشاره کرد. تغییر در خود الگو هم ممکن است در اثر مرور زمان بوجود آمده باشد.

اولین گام در بازشناسی یک الگوی تصویری، جمع‌آوری تعداد مناسبی نمونه از الگوهای مورد نظر (به‌طور مثال تصاویر اشیایی که قرار است بازشناسی شوند) است. این بخش زمان زیادی از فرآیند طراحی سیستم بازشناسی الگو را به‌خود اختصاص می‌دهد و گاهی اوقات با مشکلاتی همراه است. پس از جمع‌آوری نمونه‌های لازم، باید اقدام به انتخاب ویژگی[۳] کرد. انتخاب نوع ویژگی نیازمند دانش اولیه در مورد الگوها است. توانمندی ویژگی‌ها برای جداسازی نمونه‌های کلاس‌های مختلف، معیار انتخاب آن است.

پس از تعیین نوع ویژگی، باید روش یادگیری را انتخاب کرد. روش یادگیری می‌تواند از نوع بدون نظارت[۴]، با نظارت[۵] و یا ترکیبی باشد. در یادگیری با نظارت، هر الگو از مجموعه داده، با یک برچسب کلاسی همراه است. هدف این است که بر اساس نمونه‌های موجود، مدل طبقه‌بندی‌کننده را طوری بسازیم که بتواند نمونه‌هایی را که تا کنون ندیده است با کم‌ترین خطا در کلاس مربوط به خودشان دسته‌بندی کند. در یادگیری بدون نظارت، الگوها برچسب کلاسی ندارند و بر اساس شباهتشان در دسته‌های یکسان قرار می‌گیرند.

---

[1] Pixels (Picture Element)
[2] Intensity
[3] Feature
[4] Unsupervised
[5] Supervised

مرحله بعدی، استخراج ویژگی است. استخراج ویژگی به‌معنی به‌دست آوردن خصوصیات الگوهاست، به‌نحوی که این خصوصیات وجه مشترک الگوهای متعلق به یک کلاس و وجه تمایز الگوهای کلاس‌های مختلف باشند. استخراج ویژگی یک نگاشت از فضای الگوها به فضای ویژگی‌ها است و نتیجه این است که خصوصیات الگو را به‌صورت معنی‌دارتر و مناسب‌تری برای مرحله طبقه‌بندی آماده می‌کند. هر کمیت اندازه‌گیری شده، یک ویژگی از شی موردنظر را بیان می‌کند. این کار در عمل یک فرآیند رمزگذاری است. ویژگی‌هایی که از الگوها استخراج می‌شوند، دارای اهمیت یکسان نیستند و برخی از آنها نسبت به بقیه مهم‌تراند. در مرحله انتخاب ویژگی، با انتخاب ویژگی‌هایی که بیشترین اهمیت را در جداسازی کلاس‌های الگو دارند، هم بُعد بردار ویژگی کاهش داده می‌شود و هم کیفیت توصیف الگو توسط این ویژگی‌ها بهتر می‌شود.

انتخاب نوع مدل طبقه‌بندی‌کننده[۱]، مرحله بعدی در بازشناسی الگو است. طبقه‌بندی‌کننده به‌عنوان هسته اصلی یک سیستم بازشناسی الگو مطرح است. طبقه‌بندی‌کننده، هر الگوی ناشناخته را بر اساس ویژگی‌های آن، به‌یکی از کلاس‌های شناخته شده نسبت می‌دهد.

یک طبقه‌بندی‌کننده، میزان تعلق بردار ویژگی $x \in \Re$ به‌هر یک از کلاس‌های الگو را به‌صورت یک عدد حقیقی بیان می‌کند. مدل‌های طبقه‌بندی‌کننده‌ی متعددی برای بازشناسی الگو پیشنهاد شده‌اند که برای مثال می‌توان به شبکه‌های عصبی[۲]، طبقه‌بندی‌کننده SVM[۳] و طبقه‌بندی‌کننده کمترین فاصله[۴] و یا نزدیک‌ترین همسایه[۵] اشاره کرد.

پس از انتخاب مدل طبقه‌بندی‌کننده، باید پارامترهای آنرا مشخص کرد. روش‌های مختلفی برای تعیین این پارامترها ارایه شده‌اند. گروهی از این روش‌ها بر اساس خواص آماری، مانند بردار میانگین و ماتریس کواریانس نمونه‌های آموزشی هر کلاس (مانند آنچه که در طبقه‌بندی‌کننده بیز[۶] می‌باشد) پارامترهای طبقه‌بندی‌کننده را تعیین می‌کنند. همچنین گروه دیگری از روش‌ها وجود دارند که با استفاده از الگوریتم‌های یادگیری تکراری[۷]، مانند آموزش یک طبقه‌بندی‌کننده شبکه عصبی به‌وسیله نمونه‌های یادگیری[۸]، پارامترهای طبقه‌بندی کننده را تعیین می‌کنند. در روش‌های گروه دوم، پارامترهای طبقه‌بندی‌کننده در طول فرآیند یادگیری، تعیین می‌شوند. پس از آنکه مدل طبقه‌بندی‌کننده کامل شد، با استفاده از نمونه-

[1] Classifier
[2] Neural Networks
[3] Support Vector Machine(SVM)
[4] Minimum Distance Classifier
[5] Nearest Neighbors
[6] Bayesian Classifier
[7] Iterative Algorithm
[8] Training Data (sample)

های آزمایشی، سیستم بازشناسی الگو اعتبارسنجی[۱] می‌شود. اگر نتیجه اعتبارسنجی با مقایسه با معیارهایی که برای هر مدل متفاوت است، رضایت بخش نبود، باید یک یا چند مورد از کارهای زیر را انجام داد:

الف- ویژگی‌های استخراج شده از الگو یا ویژگی‌های انتخاب شده از بین ویژگی‌های قبلی را بهبود دهیم.

ب- مدل طبقه‌بندی‌کننده را عوض کرده و یا پارامترهای آنرا تغییر می‌دهیم.

ج- یادگیری (تعلیم) طبقه‌بندی‌کننده را کامل‌تر کرد.

کارهای فوق آنقدر تکرار می‌شوند تا تضمین کافی جهت یک تعمیم معتبر حاصل شود.

در یادگیری بدون نظارت، پس از جمع آوری داده‌های کافی از الگوهای مورد نظر، یک روش خوشه‌بندی[۲] انتخاب می‌شود (به‌عنوان مثال روش K میانگین[۳]) و بر اساس آن، داده‌ها خوشه‌بندی می‌شوند. اگرنتیجه خوشه‌بندی قابل قبول نباشد (البته می‌توان معیارهای متفاوتی را برای رد و قبول خوشه‌ها معرفی کرد)، با تغییر روش خوشه‌بندی یا تغییر پارامترهای خوشه‌بندی، کار ادامه پیدا می‌کند تا نتیجه مورد نظر حاصل شود. بازشناسی اشیا بالاترین پله مسأله بینائی سلسله مراتبی[۴] است.

۱-۳- مراحل محاسباتی درک تصویر

مراحل محاسباتی درک تصاویر را می‌توان به‌طور کلی به سه بخش تقسیم نمود:

۱ـ بینایی اولیه[۵]

۲ـ بینایی میانی[۶]

۳ـ بینایی سطح بالا[۷]

---

[1] Validation
[2] Clustering
[3] K-Means
[4] Hierarchical
[5] Primary Vision
[6] Intermediate Vision
[7] High Level Vision

در بینایی اولیه، تصویر برای استخراج ساختارهای مفید دو بعدی پردازش می‌شود. لایه میانی، از خروجی مرحله قبل برای استخراج پارامترهای مهم از تصاویر و ویژگی‌های شی استفاده می‌کند. بینایی سطح بالا نیز خروجی لایه میانی را برای انتصاب الگو به‌یکی از کلاس‌های موجود، مورد استفاده قرار می‌دهد. در ادامه به تشریح هر یک از این مراحل می‌پردازیم.

۱-۳-۱- بینایی اولیه

هدف اصلی در این لایه درک قانون‌مندی‌های تصویر می‌باشد. این قوانین ممکن است به‌صورت بافت تصویر[1]، لبه، شیب و یا تقارن باشد. برای هر کاربرد خاص جدا کردن قسمت مورد نظر از تصویر، یکی از مهم‌ترین و در عین حال مشکل-ترین بخش‌های سیستم است. خروجی پردازش در لایه اول بخش‌بندی تصویر[2] به تکه‌هایی است که هر کدام از آنها نیاز به تفسیر دارند. به‌عنوان مثال در مورد تصاویر چهره، خروجی لایه اول جداسازی محل صورت از زمینه و یا بقیه تصویر می‌تواند باشد.

۱-۳-۲- بینایی لایه‌ی میانی

به‌طور کلی قواعد اولیه‌ی تصویر به‌صورت مستقیم به توصیف معنایی اشیا ارتباط ندارند. هدف از لایه میانی رفع این مشکل و بازیابی نمایشی از تصویر با ساختار دو بعدی تصویر است. در مورد بازشناسی اشیا، اولین روشی که به‌ذهن می‌رسد، مقایسه تصویر ورودی با تصاویر موجود در بانک اطلاعاتی مدل است. روشن است که این روش کارآیی لازم را نخواهد داشت. زیرا در ابتدا این روش کاملا به تغییرات کوچک در اندازه حساس است و اگر تصویر ورودی کمی کوچک‌تر و یا بزرگ‌تر از تصویر اصلی باشد، اختلاف بین آنها زیاد خواهد شد. و همچنین با تغییر حالت شی (مثلا چهره) یا چرخش آن حول محور قایم یا در صفحه‌ی تصویر، تغییر در شدت روشنایی و یا سایر تغییرات، شناسایی با مشکل مواجه خواهد شد. از طرفی در این روش چون تصاویر به‌طور مستقیم با یکدیگر مقایسه می‌شوند، باید تصاویر به‌صورت کامل (یعنی تعداد زیادی از تصاویر که

[1] Image Texture
[2] Image Segmentation

تغییرات لازم را در خود داشته باشند) در بانک اطلاعاتی ذخیره شوند. این کار عملا احتیاج به فضای حافظه زیادی خواهد داشت که در کاربردهای تجاری امکان‌پذیر نمی‌باشد. علاوه بر آن، خواندن و مقایسه تمام تصاویر موجود در بانک اطلاعاتی با تصویر ورودی با استفاده از کامپیوتر، به زمان زیادی نیاز دارد. روشی که برای حل این مشکل ارایه شده است، استخراج چند متغیر مهم از تصویرهای مرجع و ذخیره و مقایسه‌ی این متغیر با مقادیری است که از تصویر ورودی استخراج شده‌اند. به پارامترهای استخراج شده، ویژگی گفته می‌شود.

خروجی لایه‌ی میانی مجموعه‌ای از ویژگی‌ها است که از تصویر اولیه استخراج شده‌اند. این ویژگی‌ها بسته به نوع کاربرد و همچنین ساختار شی که مورد بررسی قرار گرفته است تعریف می‌شوند. در هنگام تعریف نمودن این ویژگی‌ها، باید دقت زیادی به‌خرج داد. این ویژگی‌ها باید حاوی اطلاعاتی در مورد ساختار اصلی شی بوده و برای نمونه‌های مختلف از یک کلاس تا حد امکان ثابت باشند و همچنین به‌ازای کلاس‌های متفاوت با یکدیگر اختلاف داشته باشند. تعریف این ویژگی‌ها کاملا وابسته به کاربرد است، اما می‌توان خواص زیر را به‌عنوان معیاری برای ارزیابی ویژگی‌ها در نظر گرفت:

- ✓ نسبت به نویز حساس نباشند[۱].
- ✓ مستقل از مکان باشند[۲].
- ✓ مستقل از چرخش باشند[۳].
- ✓ مستقل از اندازه باشند[۴].
- ✓ برای تمام نمونه‌های ورودی قابل محاسبه باشند.
- ✓ تغییرات آنها به‌ازای کلاس‌های مختلف، زیاد باشد.
- ✓ مستقل از ویژگی‌های دیگر باشند.
- ✓ مستقل از سطح روشنایی محیط باشند.
- ✓ تغییرات آنها به‌ازای نمونه‌های مختلف یک کلاس کم باشد.

[1] Noise Immunity
[2] Shift Invariant
[3] Rotation Invariant
[4] Scale Invariant

۱-۳-۳- بینایی سطح بالا

نتیجه‌ی مرحله قبل به‌صورت مجموعه‌ای از ویژگی‌های کمّی که شی را به‌صورت خاصی نمایش می‌دهد، به این مرحله وارد می‌شود. اهداف این لایه عبارتند از:

۱ـ کم کردن از میزان داده‌ها در مورد هر شی در حالی که اطلاعات اصلی و مهم حفظ شوند (کاهش ُبعد).

۲ـ طبقه‌بندی اطلاعات باقی‌مانده.

برخلاف لایه‌ی قبل که کاملا بستگی به کاربرد خاصی دارد که برای آن طراحی می‌شود، این لایه معمولا برای اکثر کاربردها به‌صورت مشترک قابل طراحی است.

۱-۴- بازشناسی شی

بازشناسی اشیا پله آخر سلسله مراتب بینایی می‌باشد. ساختار کلی آن به‌صورت یک مسأله بسیار مشکل محاسباتی است و عجیب نیست که این امر همیشه مشکل بوده است و دروازه‌ای برای پاسخ به سوالات علوم شناختی و عصبی می‌باشد.

یک شی را می‌توان از نظر اختصاصی بودن در سطوح مختلفی قرار داد. به‌عنوان مثال یک گربه می‌تواند به‌عنوان " گربه من" شناخته شود و یا در سطح دیگری به عنوان "گربه" و باز هم در حوزه گسترده‌تر به‌عنوان "جانور پستاندار"و ... شناخته شود. بنابراین یک شی می‌تواند در سطوح خاص و متمایز بازشناسی شود پس خواهیم داشت:

- ✓ یک شی می‌تواند به‌عنوان یک عضو از یک کلاس کلی (همانند چهره یا ماشین) دسته‌بندی[1] شود.
- ✓ یک شی می‌تواند همچنین به‌صورت یک وجود منحصر به فرد همانند "گربه" یا "ماشین من" مشخص (تعیین هویت[2]) شود.

[1] Categorization
[2] Identification