بسم الله الرحمن الرحيم

مرکز تهران

پایان‌نامه برای دریافت درجه کارشناسی ارشد

در رشته مهندسی کامپیوتر – نرم افزار

دانشکده فنی و مهندسی

گروه علمی مهندسی کامپیوتر و فناوری اطلاعات

عنوان پایان‌نامه:

بهبود بازخورد ارتباط درسیستم‌های بازیابی معنایی تصاویر

با استفاده ازشبکه‌های عصبی فازی

نگارش:

شهروزنعمتی پور

استاد راهنما:

آقای دکتر جمشید شنبه زاده

استاد مشاور:

آقای دکتر رضا عسگری مقدم

مهر۱۳۹۰

سپاس از آقایان دکتر شنبه‌زاده و عسگری‌مقدم

که مرا در نگارش این پایان‌نامه یاری کردند.

تقدیم به مادر و پدر عزیزم که بعد از خدا هرچه دارم از آنهاست.

چکیده

امروزه ما در جهان ترابایت و میکروثانیه زندگی می‌کنیم، جایی که روزانه انبوهی از داده‌ها تولید میشود و در عین حال بازیابی سریع و بهینه اطلاعات مورد درخواست جوامع اطلاعاتی است. داده‌های چندرسانه‌ای شامل تصویر، صدا و ویدئو هستند. در نتیجه نیاز فوری برای ضبط، ذخیره و بازیابی چنین داده‌ای وجود دارد. تصویر ابتدائی‌ترین فرم داده چندرسانه‌ای است؛ به همین دلیل تحقیقات زیادی برای بازیابی چنین داده‌ای صورت گرفته است. سیستم‌های بازیابی معنایی تصاویر سیستم‌هایی هستند که برای پاسخگویی به این نیاز توسعه یافته‌اند. شکل اولیه این سیستم‌ها سیستم‌های بر پایه کلیدواژه بود که در نوع خود بی‌نظیر بود ولی معایب بسیاری داشت.

برای همین سیستم‌هایی برپایه خصوصیات بصری تصویر مورد توجه قرار گرفت. یکی از مشکلاتی که بیشتر سیستم‌های بازیابی معنایی تصاویر با آن روبرو هستند مسئله شکاف معنایی است که عبارتست از اختلاف بین خصوصیات سطح پایین تصویر و درک انسان هوشمند از کل تصویر.یکی از بهترین روش‌هایی که برای کاهش شکاف معنایی در چند سال اخیر به آن توجه شده است بازخورد ارتباط است. در مدل پیشنهادی پایان‌نامه شکاف معنایی در دو مرحله کاهش می‌یابد. از شبکه‌های شعاع مدار برای بهبود بازخورد ارتباط استفاده می‌شود با این تفاوت که به جای بردارهای ویژگی سطح پایین، از بردارهای ویژگی سطح بالا استفاده می‌شود. این بردارهای ویژگی سطح بالا توسط یک شبکه عصبی چند لایه آموزش دیده تولید می‌شوند. در واقع در ابتدا بردار ویژگی سطح پایین توسط یک شبکه عصبی سه لایه به بردار ویژگی سطح بالا تبدیل می‌شود. سپس این بردار ویژگی سطح بالا به عنوان ورودی شبکه RBF برای بهبود بازخورد ارتباط استفاده می‌شود. شبیه سازی الگوریتم توسط پایگاه داده COREL انجام شد. نتایج تحقیق بهبود حدود ۱۰درصد در دقت بازیابی را نسبت به روش معمول و بدون پیش پردازش نشان می‌دهد.بهبود الگوریتم برای تصاویر عمومی چندان مطلوب نبود در حالیکه برای تصاویر حاوی اشیا این بهبود قابل ملاحظه بود.

کلمات کلیدی: سیستم‌های بازیابی معنایی تصاویر، بازخورد ارتباط، شبکه‌های عصبی چندلایه، شبکه‌های عصبی شعاع‌مدار.

فهرست مطا لب

عنوان **صفحه**

فصل اول :مقدمه وطرح تحقیق

۱-۱-حوزه تحقیق..... ٤
۱-۱-۱-شرکت‌های دارای پایگاه داده‌های بزرگ..... ٤
۱-۱-۲-پزشکان برای انتقال دانش پزشکی..... ٤
۱-۱-۳-موتورهای جستجوی تصاویر،برای افزایش دقت دربازیابی تصاویر..... ٥
۱-۲-انگیزه واهمیت موضوع..... ٦
۱-۳-اهداف..... ۷
۱-٤-رویکردتحقیق..... ۸
۱-٥-نوآوری ۸
۱-٦-مروری برپایان نامه..... ۹

فصل دوم :تئوری واطلاعات پس زمینه

۲-۱-مقدمه..... ۱۲
۲-۲-قسمت‌های مختلف یک سیستمCBIR..... ۱۳
۲-۲-۱- استخراج ویژگی..... ۱۳
۲-۲-۱-۱- رنگ..... ۱٤
۲-۲-۱-۲- شکل..... ۱٦
۲-۲-۱-۳- روش‌های تحلیل تصویر..... ۲۰
۲-۲-۱-٤- استخراج ویژگی از طریق تطبیق شکل..... ۲۱
۲-۲-۱-٥-آستانه‌گیری..... ۲۲
۲-۲-۱-٦- تطبیق الگو..... ۲۲
۲-۲-۱-۷- بحث پیرامون تطبیق الگو..... ۲۳
۲-۲-۱-۸- تبدیل‌هاف..... ۲٤
۲-۲-۱-۹- تجزیهٔ فضای پارامتر..... ۲٥

عنوان **صفحه**

۲-۲-۱-۱۰-تبدیل‌هاف عمومی ۲۶
۲-۲-۱-۱۱-صورت‌های دیگر تبدیل‌هاف ۲۶
۲-۲-۱-۱۲-الگوهای شکل پذیر ۲۷
۲-۲-۱-۱۳-تکنیک مار ۲۷
۲-۲-۱-۱۳-۱-استفاده از الگوریتم حریصانه برای تکنیک مار ۲۸
۲-۲-۱-۱٤- خطوط هندسی فعال ۲۹
۲-۲-۱-۱۵-اسکلت کردن شی ۲۹
۲-۲-۱-۱٦-مدل‌های شکلی منعطف ۳۰
۲-۲-۱-۱۷- بافت ۳۱
۲-۲-۱-۱۷-۱- طبقه بندی بافت ۳۱
۲-۲-۱-۱۷-۲- ماتریس GRAY-LEVEL CO-OCCURRENCE ۳۲
۲-۲-۱-۱۷-۳- ارتباط مکانی بین المان‌های اولیه بافت ۳۳
۲-۲-۱-۱۷-٤- روش‌های ساخت یافته ۳٤
۲-۲-۱-۱۷-۵-روش‌های آماری ۳٤
۲-۲-۱-۱۷-٦- دسته بندی بافت ۳۵
۲-۲-۱-۱۷-۷- بخش بندی ۳۵
۲-۲-۲- توصیف شی ۳۵
۲-۲-۲-۱- توصیف مرز ۳٦
۲-۲-۲-۲- مفسر فوریه ۳۷
۲-۲-۲-۳- مفسر منطقه‌ای ۳۷
۲-۲-۲-٤- تئوری درک GESTALT ٤۰
۲-۲-۳-سنجه شباهت ٤۰
۲-۲-۳-۱- سنجه شباهت فازی در سیستم‌های بازیابی معنایی تصاویر ٤۱
۲-۲-٤- اندیس گذاری ٤۲
۲-۳-کدگذاری MPEG7 ..... ٤٤
۲-٤-خلاصه ٤۵

عنوان **صفحه**

فصل سوم : مروری بر تحقیقات گذشته

۳-۱-مقدمه ٤٨
۳-۲- شکاف معنایی ٤٨
۳-۳- طبقه بندی و خوشه بندی ٥٦
۳-٤- استفاده از بازخوردارتباط در جستجوی تصویر در CBIR ٥٧
۳-٥- روشهای متداول در بازخورد کاربر ٥٨
۳-٥-۱- چهارچوب بیزی برای RF ٥٩
۳-٥-۲- تکنیک‌های حرکت نقطه query در بازخورد کاربر سیستم‌های CBIR ٦٢
۳-٥-۳- مطالعه روشهای بازخورد کاربر منفی ٦٣
۳-٥-٤- یک مکانیزم دو مرحله‌ای بر بازخورد کاربر ٦٤
۳-٥-٥- یک روش برپایه log برای بازخورد کاربر ٦٥
۳-٥-٦- استفاده از خوشه بندی نیمه آموزشی همراه با RF ٦٦
۳-٥-۷-مسافت وزنی آموزش برای RF در CBIR ٦٦
۳-٥-۸- یک الگوریتم فیدبک دو مرحله‌ای برای CBIR ٦٧
۳-٦-پایکاه داده‌های موجود ٦٨
۳-۷-خلاصه ۷۰

۴.فصل چهارم:متد مورداستفاده و تکنیک‌های مفید

٤-۱-مقدمه ۷۲
٤-۲-رنگ ۷۲
٤-۳- شکل ۷۳
٤-٤- توصیف تصویر ۷٥
٤-٤-۱- فاز آموزش ۷٦
٤-٤-۲-مشخص سازی سطح بالا ۷۷
٤-٤-۳-مشکلاتی که بایستی حل شود ۷۷
٤-٥-استفاده از روش فازی در شبکه‌های عصبی RBF ۷۸
٤-٥-۱-مشکلاتی که بایستی حل شود ۸۱

عنوان **صفحه**

٤-٦-خلاصه ٨٢

۵.فصل پنجم.نتایج تحقیق

۵-۱-مقدمه ٨٤

۵-۲-الگوریتم پیشنهادی ٨٥

۵-۲-۱-پیش پردازش ٨٦

۵-۲-۲-استخراج ویژگی سطح پایین ٨٦

۵-۲-۳-تبدیل خصوصیات سطح پایین به بالا ٨٨

۵-۲-٤-استفاده ازسنجه شباهت برای مقایسه امضای تصاویر ٩٠

۵-۲-۵- استفاده از روش فازی در شبکه‌های عصبی RBF در RF ٩٠

۵-۳-نتایج تحقیق ٩٣

۵-۳-۱- تجزیه وتحلیل ٩٥

۵-۳-۲-مشخصات فنی ٩٧

۵-۳-۳-پارامترهای شبیه سازی ٩٩

۵-٤-مقایسه الگوریتم ١٠٠

۵-۵-خلاصه ١٠٤

۶.فصل ششم:خلاصه ،نتیجه گیری وپیشنهادات

٦-١-آنچه می آموزیم ١٠٨

٦-٢-محدودیت‌ها ١٠٩

٦-٣-راهکارهای پیشنهادی ١٠٩

٦-٤-تشکرات ١١٠

فهرست منا بع

واژه‌نامه انگلیسی به فارسی

چکیده‌ی انگلیسی

فهرست شکل‌ها

عنوان **صفحه**

شکل۱-۱-یک سیستم CBIR معمول ۳
شکل۱-۲-جستجوی تصویرگوگل ٦
شکل۲-۱-معماری یک سیستم CBIR ۱۳
شکل۲-۲-تصویر سمت چپ،تصویر نمونه، سمت راست الگوی موردنظر ۲۳
شکل۲-۳-نتیجه اعمال الگوریتم روی تصویرنمونه ۲۳
شکل۲-٤-عملکردمار ۲۸
شکل۲-۵-نمونه ای ازکدزنجیره ای ۳۷
شکل۲-٦-عدد اویلردراین شکل ٤ است(۷ شی و۳ سوراخ) ۳۹
شکل۲-۷-داده‌های دوبعدی ٤۳
شکل۲-۸-درخت مربوط به شکل قبل ٤۳
شکل۲-۹-مراحل کاردریک سیستم بازیابی معنایی صاویر ٤٤
شکل۲-۱۰-انواع امضای تصاویر ٤٤
شکل۳-۱-شکاف معنایی بین پرسش دوچرخه و نتیجه چرخ ٤۹
شکل۳-۲-سلسله مراتب شکاف معنایی برای تصویرگرگ ۵۰
شکل۳-۳-نرم افزارairline.net ۵۲
شکل۳-٤-انواع جستجو وکاربر ۵۳
شکل۳-۵-سلسله مراتب کاربران مختلف ۵٤
شکل۳-٦-مشکل مینیمم محلی ٦۲
شکل۳-۷-مشکل همگرایی کم ٦۲
شکل٤-۱-یک شبکه سه لایه feed-forward ۷٦
شکل٤-۲-مراحل مشخص سازی سطح بالا ۷۷
شکل٤-۳-یک شبکه RBFنمونه با سه لایه ۷۹
شکل٤-٤-نحوه کارgradient-descent ۸۰
شکل۵-۱-نرخ recallو precision ۹۵

عنوان **صفحه**

شکل۵-۲-عکس فوری از برنامه کاربردی-۱..... ۹۷
شکل۵-۳-عکس فوری از برنامه کاربردی-۲..... ۹۸
شکل۵-٤-آموزش شبکه چندلایه پرسپترون..... ۹۹

فهرست جدول‌هاو نمودا رها

عنوان **صفحه**

جدول۲-۱-مقایسه روش‌های مختلف تشخیص لبه..... ۱۹
جدول۲-۲-تکنیک‌های مورداستفاده درتشخیص شکل..... ۲۲
جدول۲-۳-تکنیک‌های تشخیص بافت..... ۳۲
جدول۲-٤-تکنیک توصیف شکل..... ۳٦
جدول۲-٥-مقایسه پیچیدگی زمانی سنجه‌های شباهت..... ٤۱
جدول۳-۱-روش‌های کاهش شکاف معنایی..... ٥۱
جدول۳-۲-انواع پایگاه تصاویر مورداستفاده درCBIR..... ٦۹
جدول٥-۱-گروه‌های معنایی پایگاه داده آموزشی..... ۹۳
جدول٥-۲-گروه‌های معنایی موجوددرپایگاه تصاویر تست..... ۹٤
جدول ٥-۳-تنظیمات پارامترهای الگوریتم..... ۱۰۰
جدول ٥-٤-دقت سیستم پیشنهادی با پیش پردازش و بدون آن..... ۱۰۲
جدول ٥-٥-دقت بازخورد ارتباط باپیش پردازش و بدون آن..... ۱۰۳
جدول ٥-٦- دقت بازخورد ارتباط باپیش پردازش و بدون آن برای تصاویر عمومی..... ۱۰٤
نمودار۲-۱-فلوچارت الگوریتم مار..... ۲۸
نمودار۲-۲-محدوده استاندارد MPEG-7..... ٤٥
نمودار۳-۱-چهارچوب بیزی برای بازخوردکاربر..... ٦۱
نمودار٥-۱-مدل پیشنهادی CBIR..... ۸٥
نمودار٥-۲-مقایسه نمودارprecision-recallبرای روش مذکور..... ۱۰۱
نمودار٥-۳-همگرایی سریع متد پیشنهادی..... ۱۰۲
نمودار ٥-٤-عدم بهبود روش پیش پردازش برای تصاویر عمومی..... ۱۰۳

۱. فصل اول:مقدمه وطرح تحقیق

شبکه بینایی انسان طی قرن‌ها تکامل پیدا کرده است و قدرت تشخیص و تفکیک فوق العاده زیادی را کسب کرده است، به طوریکه هرگونه شبیه‌سازی کامپیوتری از آنها در طول سالیان اخیر به شکست خورده است و یا حداقل دقت کافی را نداشته است. یکی از دلایلی که سیستم‌های تصویری به اندازهٔ سیستم‌های بر پایهٔ متن موفق نبوده‌اند را می‌توان به تفاوت ماهوی متن و تصویر نسبت داد. کلمه حاصل ساخت بشر است ولی تصویر از بدو شکل‌گیری بشریت همراه او بوده است.

یکی از آرزوهای انسان انتقال قدرت آنالیز تصاویر به کامپیوتر است، به طوریکه بدون دخالت انسان، تصاویر طبقه‌بندی و مدیریت شوند. همچنین طی چند سال اخیر، اطلاعات زیادی در قالب مولتی‌مدیا، در پایگاه داده‌های فعلی ذخیره شده‌اند. ایجاد روشی برای بازیابی اطلاعات به خصوص تصاویر بسیار ضروری است.

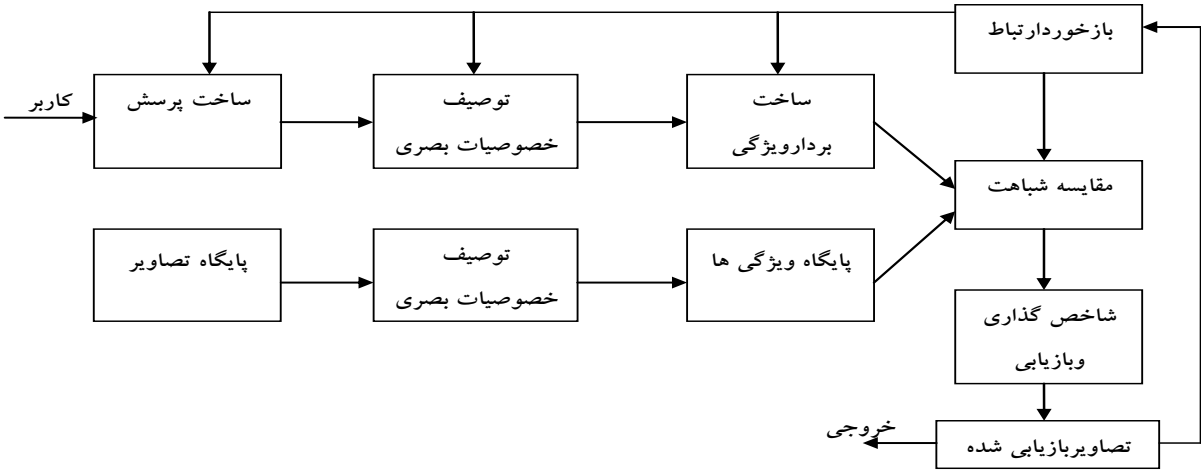
یکی از سیستم‌هایی که در این زمینه طراحی شده است، سیستم CBIR[1] است. سیستم بازیابی معنایی تصاویر (CBIR) برخلاف جد خود TBIR[2] براساس خصوصیات بصری تصویر عمل می‌کند تا ضعف TBIR را که بر پایهٔ کلمه است را بپوشاند.

طیف و دامنهٔ تحت پوشش CBIR از یک تابع ابتدایی شباهت تصویر تا یک موتور جستجوی تصویر را در بر می‌گیرد. در واقع می‌توانیم CBIR را نقطهٔ اتصال[3] بین علوم زیر بدانیم:

۱- بینایی ماشین
۲- یادگیری ماشین
۳- بازیابی اطلاعات
٤- ارتباط انسان و ماشین
٥- پایگاه داده‌ها
٦- داده کاوی

معماری یک سیستم CBIR معمولی در شکل ۱-۱ آمده است:

[1] Content-Based Image Retrieval
[2] Text-based image retrieval
[3] Juncture



شکل۱-۱- یک سیستم CBIR معمول

در این سیستم‌ها به طور معمول دو مرحله عملیات صورت می‌گیرد:

۱ -پردازش برون خطی[۱]

۲ -پردازش درون خطی[۲]

در مرحلهٔ اول تمام تصاویر در پایگاه داده پردازش می‌شوند. ویژگی‌های مستخرج متادیتا[۳] را می‌سازند. سپس این ویژگی‌ها که برای اندیس‌گذاری تصاویر به کار می‌روند همراه با تصویر در پایگاه داده ذخیره می‌شوند.

در مرحلهٔ دوم که همان مرحلهٔ بازیابی است، تصویر پرسش آنالیز می‌شود و ویژگی‌های آن نیز استخراج می‌شوند. سپس توابع خاصی که مقیاس شباهت[٤] نام دارند برای جستجو و بازیابی تصاویری که در پایگاه داده شبیه به تصویر پرسش هستند به کار می‌روند. در واقع به جای مقایسهٔ دو تصویر با یکدیگر، ویژگی‌های آنها با یکدیگر مقایسه می‌شوند. همانگونه که از توضیحات بر می‌آید، یک سیستم مدرن بازیابی معنایی تصاویر شامل قسمت‌های کلیدی زیر است:

۱ -استخراج ویژگی[۵]

۲- نمایش ویژگی[٦]

۳- فهرست گذاری[۷]

٤- سنجه شباهت[۸]

۵- بازخورد ارتباط[۹]

٦- بهینه سازی[۱]

[1] Offline processing
[2] Online processing
[3] Meta-Data
[4] Similarity Measure
[5] Feature extraction
[6] Feature representation
[7] indexing
[8] Similarity measure
[9] Relevance feedback

سیستم‌های مختلفی بر پایه جستجوی معنایی تصاویر(CBIR) در یک دهه اخیر شکل گرفته‌اند ولی بیشتر آنها نیازهای کاربر را برآورده نمی‌کنند و به صورت تجاری در نیامده‌اند. مسئله اصلی به حداکثر رساندن دقت مرتبط بودن تصاویر بازیابی شده با پرسش وارد شده از طریق عامل انسانی است.

در این زمینه سئوالات زیر مطرح هستند:

۱ –چرا نرم‌افزارهای CBIR امروزی هنوز نتوانسته‌اند جایگاه خود را پیدا کنند؟

۲ –آیا می‌توان روشی برای به حداکثر رساندن دقت تصاویر بازیابی شده پیدا کرد؟

۳ –آیا می‌توان با تست روی نمونه کوچک، قضیه رابه کل تصاویر عمومیت داد؟

٤ –آیا دلیل ناقص بودن سیستم‌های فعلی، عدم تعامل با یک کاربر و استفاده از خصوصیات سطح پایین تصویراست؟

۵ –چگونه می‌توان بازخورد کاربر را به بهترین شکل در سیستم‌های CBIR فعلی لحاظ کرد؟

۱–۱– حوزه تحقیق

۱–۱–۱– شرکت‌های دارای پایگاه داده‌های بزرگ

همانگونه که داده‌های یک شرکت حکم ارزشمندترین موجودیت را دارند، تمام شرکت‌ها علاقه زیادی برای ذخیره، دسته‌بندی و بازیابی مطلوب آنها دارند. در گذشته بیشتر اطلاعات به صورت متنی بوده است و بر مبنای آن تکنیک‌های مختلف داده‌کاوی [۲]توسعه یافته‌اند. در حال حاضر مسئله تصاویر و به طور کلی فایل‌های چندرسانه‌ای کم‌کم جای خود را در شرکت‌ها باز می‌کنند و حجم بالایی از اطلاعات به صورت تصویر تولید می‌شود. برای همین، نقش سیستم‌های جستجوی معنایی تصاویر اینجا نمود پیدا میکند چرا که با وجود تعداد تصاویر بالا، امکان توصیف تصاویر به صورت متنی وجود ندارد و در صورت امکان‌پذیر بودن، هزینه اپراتوری زیادی تحمیل می‌شود. در نتیجه مدیران شرکت‌ها خواستار سیستم‌های بهینه‌ای هستند که این عملیات را به صورت خودکار و حداقل نیمه‌خودکار انجام دهند.

۱–۱–۲– پزشکان برای انتقال دانش پزشکی

با توجه به فراگیرشدن پزشک راه دور(ملاقات با پزشک در دنیای مجازی) و در دسترس نبودن پزشکان در همه مواقع، بایستی جوابی به بیمار در خصوص بیماری داده شود. امروزه وقت بسیاری از پزشکان صرف تشخیص دادن اولیه تصاویری است که توسط مراکز رادیولوژی و سونوگرافی گرفته

[1] Optimization
[2] Data mining

می‌شود. بررسی‌ها نشان می‌دهد که بیشتر بیماران مشکل خاصی ندارند. از این رو تشخیص اتوماتیک تصاویر اعضای بدن بیماران یکی از نیازهای پزشکان در شرایط فعلی است. سیستم‌های CBIR در صورت آموزش درست می‌توانند به پزشک در تشخیص اولیه بیماری کمک کنند. با توجه به اینکه در حال حاضر پایگاه داده‌های بزرگی از تصاویر اعضای بدن انسان‌ها در بیمارستان‌ها موجود است، می‌توان با انتقال دانش اساتید علم پزشکی به این نوع سیستم‌ها و به عبارت دیگر هوشمند کردن آنها، کمک بزرگی به پزشکان در تشخیص بیماری کرد. با استفاده از این سیستم مراجعه بیمار به پزشک فقط در موارد خاص صورت می‌گیرد. لازم به ذکر است چون درصد خطای این سیستم‌ها بالای ۵۰درصد است، تا رسیدن به دقت در مرز ۹۷ درصد زمان زیادی باقیمانده است.

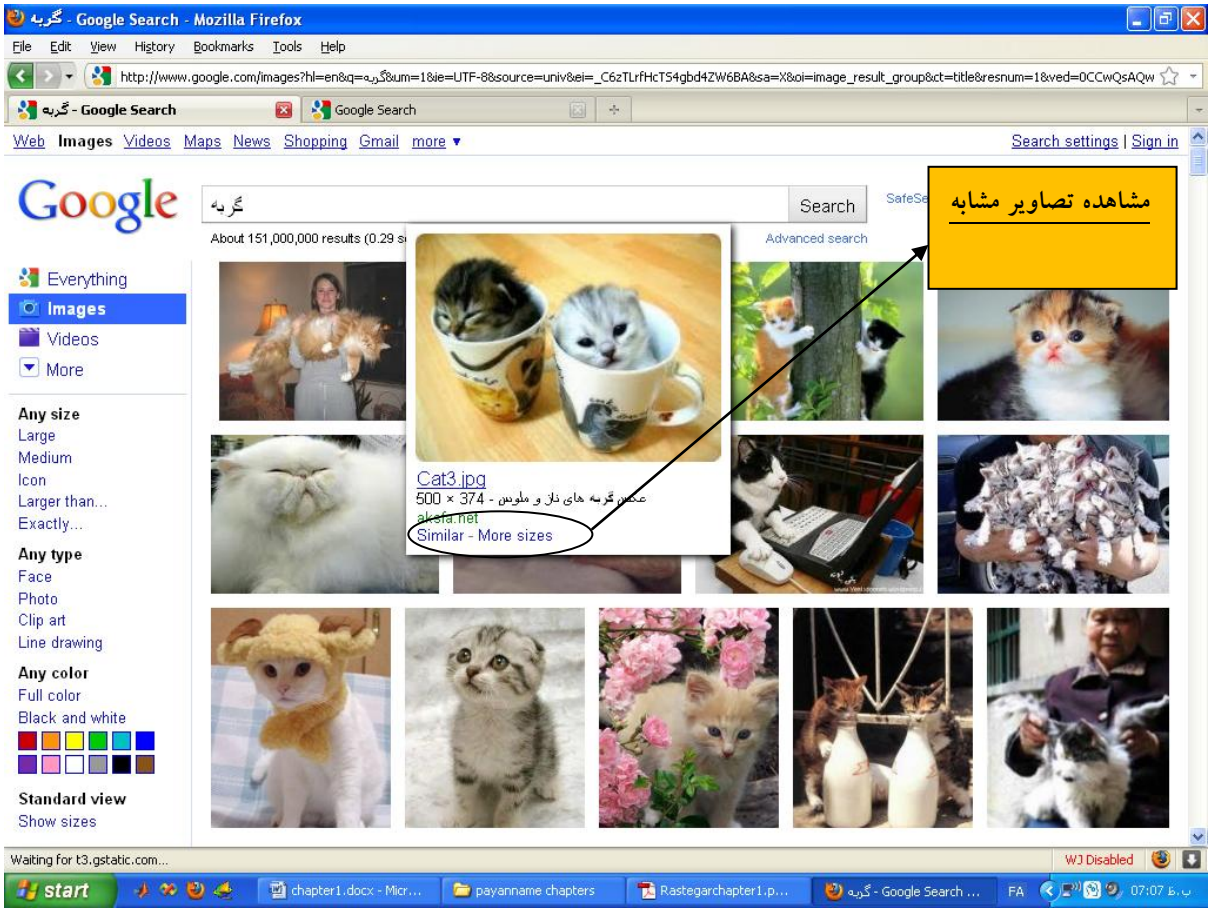
۱–۱–۳– موتورهای جستجوی تصاویر، برای افزایش دقت دربازیابی تصاویر

موتورهای جستجوی معروف مثل گوگل [1]و بینگ [2]که مدت زیادی است به خوشه‌بندی و دسته‌بندی اطلاعات متنی موجود در اینترنت مشغول هستند، اخیرا موتور جستجوی تصویر خود را که قبلا براساس زیرنویس [3]های کنار تصویر بود، بهبود داده‌اند. هنگامی که مکان‌نما روی هرتصویر برده می‌شود، امکان مشاهده تصاویر مشابه [4]وجود دارد که تصاویر مشابه به پرسش درخواستی را به ما نشان می‌دهد. در شکل ۱–۲ این قضیه با پرسش "گربه" نشان داده شده است.

همانگونه که مشاهده می‌شود، با بردن مکان نما به روی تصویر دوم، دو گزینه "similar images"و"more sizes" آشکار می‌شود. اساس کار این ویژگی همان مسئله مورد بحث یاCBIR است که موتور جستجو با مراجعه به سایت‌های مختلف، تصاویر را براساس مدل خود فهرست [5]می‌کند که با توجه به اینکه این قبیل سیستم‌ها هنوز کارایی لازم را ندارند، بهبود این قبیل سیستم‌ها برای شرکت‌های بزرگی چون گوگل نیاز حیاتی است.

علاوه بر این چون سیستم بازرگانی فعلی گوگل یا همان adwork که مربوط به تبلیغات است، صرفاً براساس متن کار می‌کند، می‌توان با افزایش دقت سیستم‌های جستجوی معنایی تصاویر، بحث تبلیغات تصویری را نیز به این سامانه اضافه کرد تا تحول عظیمی در بحث تبلیغات ایجاد کرد.

[1] Google
[2] Bing
[3] caption
[4] similar images
[5] index



شکل ۱-۲- جستجوی تصویرگوگل

۱-۲- انگیزه واهمیت موضوع

مسئله شکاف معنایی مسئله‌ای بسیار مهم در زمینه سیستم‌های CBIR است. در طول یک دهه گذشته کارهای زیادی در مورد بازیابی تصاویر براساس خصوصیات سطح پایین تصویر از جمله رنگ، بافت و شکل صورت گرفته است[۱و۵و۷].

یکی از مشکلاتی که محققان با آن روبرو شده‌اند، مسئله شکاف معنایی است. شکاف معنایی اختلاف بین درک هوشمند بشری از تصویر و خصوصیات سطح پایین تصویر است که از طریق محاسبات ریاضی به دست می‌آید [۶۰و۶۱و۶۲و۶۳].

درک انسان از یک تصویر وابسته به عوامل بسیاری از جمله دانش، احساسات و غیره است که کامپیوتر از درک آن عاجز است. گنجاندن معنا در تصاویر به بهبود عملکرد CBIR کمک می‌کند. در واقع مسئله شکاف معنایی از دو بعد قابل بررسی است: ۱- پرکردن شکاف از پایین، به این صورت که خصوصیات سطح پایین به طریقی تبدیل به خصوصیات سطح بالا شوند تا بتوانند به

صورت فشرده دارای معنا و مفهوم تصویر باشند. ۲- پرکردن شکاف از بالا، که بدون تغییر خصوصیات سطح پایین و بردار مربوط به آن، ازطریق دخیل کردن عامل انسانی، نتیجه جستجو را به سمت بهینه شدن پیش می‌بریم. یکی از روش‌هایی که برای پر کردن این شکاف به روش دوم طی سال‌های گذشته به کار گرفته شده است، مسئله دخیل کردن کاربر در پروسه استخراج معنایی تصاویر است[۳۶و۳۷]. بازخورد ارتباط[۱] طی چند سال اخیر مورد توجه بسیاری از محققان قرار گرفته است. به دلیل پروسه طولانی بازخورد ارتباط در سیستم‌های فعلی، غالبا از بازدهی آنها کاسته می‌شود، چون وقت زیادی را از کاربر می‌گیرد. تاکنون روش‌های مختلفی در این زمینه به کار رفته‌اند که مهمترین آنها به قرار زیر هستند[۳۸و۴۱و۴۶و۴۷و۴۹و۵۷]:

۱ -متدولوژی بیزی[۲]
۲ -شبکه‌های عصبی
۳ -آرایه‌های متعامد[۵]
۴- متدهای فازی[۳]
۵- الگوریتم‌های ژنتیک[۴]
۶- BMPM[۶]

روش‌های مذکور دارای معایبی از جمله دقت کم، سرعت کم و خاص منظوره بودن هستند.
ایجاد روشی برای هر چه بیشتر خودکار کردن این پروسه و دخیل کردن هوش انسانی در آن می‌تواند نقش عمده‌ای در بهبود عملکرد سیستم‌های فعلی CBIR داشته باشد.

۱-۳- اهداف

اهداف این تحقیق را می توانیم به صورت زیر خلاصه کنیم:

۱ -بررسی دلایل ضعف روش‌های موجود.
۲ -ارائه یک روش بر اساس شبکه‌های عصبی فازی برای بهبود روش‌های فعلی و ایجاد برچسب معنایی[۷] برای تصاویر هدف.
۳ -بررسی نقاط ضعف یا قوت روش مذکور از طریق مقایسه آن با روش‌های قبلی.

در واقع با توجه به عدم موفقیت شایان روش‌های فعلی(هدف اول)، شناخت نقاط ضعف و قوت این روش‌ها می‌تواند به ما در ایجاد روش جدید کمک کند. بسیاری از روش‌های قبلی با توجه به بازخوردهای متعدد، بسیار وقت گیر و از حوصله کاربر خارج بوده‌اند.

[1] Relevance feedback
[2] Bayesian methodology
[3] fuzzy methods
[4] genetic programming
[5] Orthogonal arrays
[6] Biased mixture probability machine
[7] Semantic tag

تاکنون روش‌های مختلفی بر اساس منطق فازی و شبکه‌های عصبی ارائه شده است. شبکه‌های عصبی برای کلاسه‌بندی الگوها از دیرباز مورد توجه قرار داشته‌اند. منطق فازی نیز در زمینه دقت عمل، بیشتر به دنیای واقعی شباهت دارد. بازخوردهای کاربر بیشتر به صورت تصاویر مثبت و منفی با توجه به پرسش داده شده است. این مسئله بیشتر شبیه به یک مسئله کلاسیک دسته‌بندی است. در بعضی موارد تقسیم تصویر بازگردانده شده به دو حالت مثبت و منفی نمی‌تواند ما را در رسیدن به اهداف یاری کند. در نتیجه مجموعه‌های چند حالتی(فازی) می‌توانند مفید باشند.

شبکه‌های عصبی فازی (هدف ۲) قدرت دسته بندی شبکه عصبی و دقت شبکه فازی را با یکدیگر ترکیب می‌کنند تا بازخوردهای کاربران را به بهترین نحو برای فهرست کردن تصاویر مدیریت کنند.

۱–۴– رویکرد تحقیق

رویکرد این تحقیق استفاده از شبکه عصبی فازی برای بهبود بازخورد ارتباط برای سیستم‌های بازیابی معنایی تصاویر است. در واقع در این پایان‌نامه ما در دو گام شکاف معنایی را کم می‌کنیم. گام اول استفاده از شبکه‌های عصبی چندلایه برای تبدیل بردار ویژگی سطح پایین به بردار ویژگی سطح بالاست. در گام دوم با استفاده از یک شبکه شعاع مدار تابعی، بازخوردهای کاربر را که بیشتر به صورت فازی است را مدیریت می‌کنیم و شبکه را به صورت برخط آموزش می‌دهیم تا در تکرارهای بعدی جواب‌های تا حد امکان مرتبط با پرسش داده شود.

۱–۵– نوآوری

نوآوری این پایان‌نامه، رویکرد دووجهی به قضیه شکاف معنایی است. تا به حال سیستم‌های بازخورد ارتباط با بردار ویژگی سطح پایین کار کرده‌اند و با توجه به این نکته که معمولا اندازه بردار ویژگی سطح پایین در سیستم‌های CBIR فعلی بیش از ۱۰۰ است، این سیستم‌ها خیلی دیر همگرا می‌شوند و در ضمن سرعت پایینی دارند. در نتیجه تغذیه ورودی سیستم‌های RF توسط بردار ویژگی سطح بالا، که خیلی خوب فشرده شده است، می‌تواند در بهبود عملکرد بازخورد ارتباط کمک کند. این پایان‌نامه از بردار ویژگی سطح بالا برای ورودی شبکه RBF [1] استفاده می‌کند.

می‌توانیم نوآوری پایان‌نامه را به صورت زیر خلاصه کنیم:

۱. استفاده از شبکه عصبی فازی [2] در بازیابی تصاویر.

[1] Radial basis function
[2] Fuzzy neural network