

انتخاب ویژگی جهت ارتقاء دقت سیستم های بازیابی تصویر مبتنی بر محتوا

ابوالفضل تذری
گروه هوش مصنوعی دانشگاه آزاد قزوین
Tazaree62@Yahoo.com

کریم فائز
دانشکده مهندسی برق دانشگاه صنعتی امیر کبیر
و دانشکده مهندسی برق و رایانه دانشگاه آزاد قزوین
Kfaez@aut.ac.ir

کمک بازخورد مرتبط پرداخته اند. همچنین عده ای نیز به بهبود دسته بندی و بازیابی، تواما، روی آوردند [4,5,6]. نکته ای که کمتر به آن توجه شده است، کیفیت ویژگی های مورد استفاده می باشد. ما در این مقاله مرحله جدید «انتخاب ویژگی» را برای افزایش دقت دسته بندی در سیستم های CBIR پیشنهاد می کنیم. برای اعمال الگوریتم پیشنهادی تعداد ۴۵۵ تصویر را در ده گروه مفهومی، از پایگاه داده LabelMe [7] برگزیدیم (جدول ۱). این پایگاه داده دارای تصاویر از پیش تقطیع شده می باشد. شکل ۱ چند نمونه از تصاویر مورد استفاده را نشان می دهد.

جدول ۱: گروه های مفهومی و تعداد تصویر در هر گروه

شماره	گروه مفهومی	تعداد	شماره	گروه مفهومی	تعداد
۱	درخت	۳۰	۶	ابر	۵۰
۲	ساحل	۵۰	۷	کتابخانه	۵۰
۳	آسمان	۴۵	۸	پنجره	۵۰
۴	خورشید	۴۵	۹	جاده	۴۵
۵	دست	۴۵	۱۰	دیوار	۴۵



شکل ۱: چند نمونه تصویر تقطیع شده از تصاویر پایگاه داده

۲- چارچوب سیستم

شکل ۲ مراحل الگوریتم پیشنهادی را در دو فاز آموزش و تست نشان می دهد. در فاز آموزش به ساخت دسته بندی (Classifier) و در فاز تست به بازیابی و نمایش تصاویر می پردازیم. ما سه نوع ویژگی سطح پایین رنگ، بافت و شکل را، در دو فضای رنگی RGB و HSV، بطور محلی

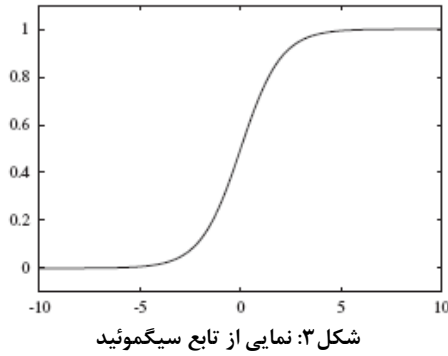
چکیده: در این مقاله روشی جهت افزایش دقت سیستم های بازیابی تصویر مبتنی بر محتوا (CBIR: Content Based Image Retrieval) ارائه می شود. این سیستم ها ابتدا به دسته بندی تصاویر پرداخته و سپس به بازیابی دسته ای از تصاویر که مورد نظر کاربر است اقدام می کنند. ما با انتخاب ویژگی های کیفی، به ارتقاء دقت دسته بندی می پردازیم. در روش پیشنهادی قدرت دسته بندی هر ویژگی به کمک یک معیار آماری محاسبه شده و سپس به گزینش ویژگی هایی می پردازیم که قدرت دسته بندی بیشتری داشته باشند. نتایج آزمایشی نشان می دهند، تعداد کمی از این ویژگی ها قادرند پیچیدگی محاسباتی را کاهش و دقت دسته بندی را افزایش دهند.

واژه های کلیدی: بازیابی تصویر مبتنی بر محتوا، بازشناسی الگو، هوش محاسباتی، انتخاب ویژگی، دسته بندی تصاویر، شبکه های عصبی مصنوعی.

۱- مقدمه

افزایش روز افزون حجم عظیم اطلاعات چند رسانه ای، «مدیریت محتوا» را به عنوان یک نیاز راهبردی مطرح می سازد. «بازیابی تصویر مبتنی بر محتوا» یکی از رهیافت های مدیریت محتوای بصری می باشد. این رهیافت شامل تکنیک هایی است که به دسته بندی و بازیابی تصاویر یک پایگاه داده می پردازند [1]. فرض اساسی سیستم های CBIR آن است که تصاویری که دارای مفاهیم مشابهی هستند، دارای ویژگی های بصری مشابهی نیز می باشند. بنابراین این سیستم ها لزوماً قادر به درک محتویات تصاویر نیستند بلکه به کمک ویژگی های بصری، تصاویر را به گروه هایی با مفاهیم مشابه دسته بندی می کنند [1]. سیستم های CBIR، علاوه بر «دسته بندی»، از تکنیک «بازخورد مرتبط» (Relevance Feedback) جهت ارتقاء دقت بازیابی استفاده می کنند [2].

با ارتقاء کیفی هر یک از دو مرحله فوق (دسته بندی و بازیابی) دقت نهایی سیستم های CBIR افزایش می یابد. برخی محققان به افزایش دقت دسته بندی [1,3] و برخی دیگر به افزایش دقت بازیابی به



۴- گسسته سازی

هدف از گسسته سازی بردار ویژگی، کاهش حجم داده های ورودی است، به گونه ای که از دقت دسته بندی کاسته نشود [3].

در این بخش روش مورد استفاده در گسسته سازی را تشریح می کنیم. مجموعه مقادیر ویژگی f_k در بردارهای ویژگی را با $f_k = \{x_1, x_2, \dots, x_n\}$ نمایش می دهیم. که در آن x_i بیانگر مقدار ویژگی f_k در i -امین بردار ویژگی می باشد. بدون از دست دادن عمومیت مساله، بردارهای ویژگی را بگونه ای مرتب می کنیم که برای مجموعه مقادیر ویژگی f_k داشته باشیم:

$$x_1 \leq x_2 \leq \dots \leq x_n$$

اکنون مجموعه x_i ها را از نقطه x_p به دو بخش تقسیم می کنیم. مقدار مناسب x_p برای تقسیم ویژگی f_k در مجموعه بردارهای ویژگی U ، مقداری است که بتواند E را کمینه سازد [9]:

$$E(f, x_p) = \frac{|S_1|}{|U|} Entropy(S_1) + \frac{|S_2|}{|U|} Entropy(S_2) \quad (2)$$

که در آن:

$$Entropy(S) = - \sum_{c \in C} \frac{|Sc'|}{|S|} \log_2 \frac{|Sc'|}{|S|} \quad (3)$$

U : مجموعه بردارهای ویژگی

S_1 : مجموعه بردارهای ویژگی که در آن: $f_k \leq x_p$

S_2 : مجموعه بردارهای ویژگی که در آن: $f_k > x_p$

C : مجموعه مفاهیم مورد بررسی - مفاهیم دهگانه جدول (۱)

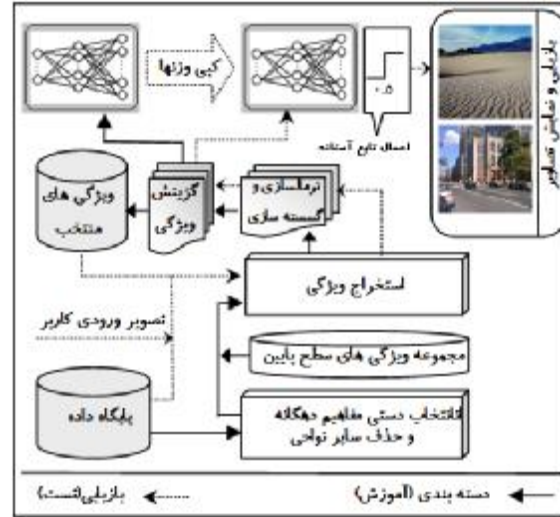
Sc' : مجموعه بردارهای ویژگی که حاوی مفهوم c هستند.

یافتن x_p و تقسیم مجموعه داده ای را تا برقراری یکی از دو شرط زیر ادامه می دهیم:

۱- زیر مجموعه ایجاد شده خالص گردد (بردارهای ویژگی در آن زیر مجموعه تنها شامل یک مفهوم باشند)

۲- برای تمام ویژگی ها، x_p مناسب را پیدا کرده باشیم.

هنگامی که یکی از دو شرط فوق برقرار شد، بردارهای ویژگی به زیر مجموعه هایی، که هر یک بیشترین خلوص را دارند، تقسیم می شوند.



شکل ۲: چارچوب سیستم

از هر ناحیه تصویر استخراج نمودیم. هر یک از این ویژگی ها دارای خصیصه هایی می باشند. در جدول ۲ برخی از خصایص این ویژگی ها را ملاحظه می فرمایید.

جدول ۲: برخی خصیصه های استخراج شده در بردار ویژگی

ویژگی	خصیصه
رنگ	میلتگین، میانه، واریانس، ممان های سوم تا پنجم، و ...
بافت	انرژی، تمایز، همبستگی، همبستگی در چهار جهت اصلی از همسایگی اول تا پنجم
شکل	گشتاورهای سوم تا پنجم

۳- نرمالسازی بردار ویژگی

منظور از نرمالسازی توزیع مقادیر ویژگی در بازه ای از قبل تعیین شده، اغلب بین صفر و یک، می باشد [8]. روش مورد استفاده ما برای نرمالسازی ویژگی i از بردار k عبارتست از:

$$x'_{i,k} = \frac{x_{i,k}}{\max(x_{i,1}, \dots, x_{i,n})} \quad (1)$$

ضرورت اعمال نرمالسازی از آن جهت است که مشتق تابع فعالیت شبکه عصبی به عنوان مضربی در معادله تغییر وزن ظاهر می شود. لذا اگر مشتق تابع فعالیت به صفر میل کند، میزان تغییر وزن های پیشین شبکه نیز به صفر میل خواهد کرد و عملاً شبکه آموزش نخواهد دید. اکنون این پرسش مطرح می گردد که چه هنگام مشتق تابع فعالیت به صفر میل می کند؟ همانطور که در شکل ۳ پیداست، اگر تابع فعالیت یک تابع سیگموئید باشد، مشتق آن خارج از بازه $[-5, +5]$ به صفر میل می کند. برای جلوگیری از این مشکل، باید از ورود مقادیر عددی بزرگ (خارج از بازه $[-5, +5]$) به تابع فعالیت جلوگیری نمود. ما با نرمالسازی بردار ویژگی در بازه $[0, 1]$ از بروز مشکل عدم آموزش جلوگیری نمودیم.

روش مورد استفاده در انتخاب ویژگی، استفاده از یک معیار آماری به نام «بهره اطلاعات» (Information Gain) می باشد. بهره اطلاعات، قدرت دسته بندی یک ویژگی را، در مجموعه داده ای بیان می کند [9]. هدف از دسته بندی آن است که مجموعه ورودی به دسته هایی تقسیم گردد که هر دسته تنها معرف یک مفهوم باشد. به عبارت دیگر خلوص هر دسته بیشینه باشد. از آنجاییکه ویژگی های دارای بهره اطلاعات بیشتر، مجموعه داده ای را به دو قسمت با حداکثر خلوص تقسیم می کنند، بنابراین این ویژگی ها برای دسته بندی ویژگی های مناسبی به شمار می روند [8]. «بهره اطلاعات» ویژگی A در مجموعه بردارهای ویژگی S که دارای تعداد V مفهوم می باشند عبارتست از [9]:

$$Gain(S, A) = Entropy(S) - \sum_{v \in Value(A)} \frac{|S_v|}{|S|} Entropy(S_v) \quad (4)$$

که در آن:

$Entropy(S)$: در قسمت ۴ توضیح داده شد

$Value(A)$: مجموعه مقادیر ویژگی A

S_v : زیر مجموعه ای از S که ویژگی A در آن زیر مجموعه دارای مقدار v باشد. به عبارت دیگر:

$$S_v = \{s \in S \mid A(s) = v\}$$

ما میزان بهره اطلاعات را برای تمام ویژگی های استخراج شده محاسبه کرده سپس بنابر دقت مورد نیاز، قدرت محاسباتی موجود و زمان مورد نظر برای آموزش، تعدادی از بهترین ویژگی ها را انتخاب، و بوسیله شبکه عصبی به دسته بندی تصاویر پرداختیم.

۶- نتایج آزمایشی

روش K-fold cross validation با $k=10$ در آموزش شبکه مورد استفاده قرار گرفت [9]. تعداد نرون های لایه میانی ۲۰، نرخ آموزش ۰.۳ و ممنوم برابر ۰.۲ در نظر گرفته شده است.

بررسی تاثیر مراحل سه گانه الگوریتم پیشنهادی در افزایش دقت دسته بندی را در شکل ۴ ملاحظه می فرمایید. چنانچه مشاهده می شود، هر یک از این سه مرحله، در افزایش دقت و کاهش زمان دسته بندی (تعداد ایپوک های کمتر) موثر می باشند. داده های خام در ۳۰ ایپوک، و داده های گسسته شده در ۵ ایپوک، تقریباً، به حداکثر دقت دسته بندی خواهند رسید.

سومین مرحله از الگوریتم پیشنهادی «گزینش ویژگی» می باشد. برای ارزیابی این مرحله آزمایشی زیر را انجام دادیم: پس از استخراج ویژگی و نرمالسازی و گسسته سازی بردار ویژگی، بهره اطلاعات تمام ویژگی ها را محاسبه کرده و ویژگی ها به ترتیب افزایش بهره اطلاعات مرتب نمودیم.

مثال ۱. فرآیند گسسته سازی را بر روی شش بردار ویژگی که مشتمل بر سه مفهوم می باشند، اعمال می کنیم:

مفاهیم = {کتابخانه، ساحل، درخت}

بردار ویژگی، که دارای ۵ مولفه می باشد، را در بازه [0, 1] نرمال کرده (جدول ۳-الف)، و الگوریتم گسسته سازی را اعمال می کنیم (جدول ۳-ب). نقاط برش، x_p ، برای ویژگی های اول تا پنجم به ترتیب عبارتند از: ۰.۴۵؛ ۰.۶؛ ۰.۷؛ ۰.۵۵؛ ۰.۴ یعنی:

$$\begin{array}{ll} a_1 \leq 0.45 & \text{و} & a_2 > 0.45 \\ b_1 \leq 0.6 & \text{و} & b_2 > 0.6 \\ c_1 \leq 0.7 & \text{و} & c_2 > 0.7 \\ d_1 \leq 0.55 & \text{و} & d_2 > 0.55 \\ e_1 \leq 0.4 & \text{و} & e_2 > 0.4 \end{array}$$

همانطور که ملاحظه می شود، داده های خام در قسمت (الف) دارای ۳۰ مقدار مختلف و داده های گسسته شده در قسمت (ب) تنها دارای ۱۰ مقدار می باشند. این در حالیست که دقت دسته بندی داده های قسمت (ب) با قسمت (الف) برابر می باشد.

جدول ۳: گسسته سازی بردار ویژگی

مفهوم	بردار ویژگی				
(الف) بردارهای ویژگی نرمال شده					
ساحل	۰.۲۰	۰.۸۲	۰.۳۳	۰.۷۹	۰.۶۳
کتابخانه	۰.۶۵	۰.۷۳	۰.۸۹	۰.۷۲	۰.۵۸
کتابخانه	۰.۷۷	۰.۳۴	۰.۴۲	۰.۸۱	۰.۱۳
درخت	۰.۸۰	۰.۸	۰.۷۵	۰.۲۳	۰.۱۹
ساحل	۰.۱۷	۰.۳۹	۰.۳۶	۰.۷۵	۰.۲۱
درخت	۰.۲۵	۰.۷۷	۰.۸۲	۰.۳۲	۰.۲۴
(ب) بردارهای ویژگی، پس از گسسته سازی					
ساحل	a_1	b_2	c_1	d_2	e_2
کتابخانه	a_2	b_2	c_2	d_2	e_2
کتابخانه	a_2	b_1	c_1	d_2	e_1
درخت	a_2	b_2	c_2	d_1	e_1
ساحل	a_1	b_1	c_1	d_2	e_1
درخت	a_1	b_2	c_2	d_1	e_1

۵- گزینش ویژگی

ویژگی های زیادی وجود دارند که همه آنها برای دسته بندی مناسب نیستند. به عنوان مثال Brightness برای دسته بندی ویژگی مناسبی نیست، زیرا تصاویر با روشنایی کم یا زیاد می توانند دارای هر مفهومی باشند. هدف از انتخاب ویژگی، گزینش ویژگی هایی است که توانایی برجسته ای در دسته بندی داشته باشند [8].

جدول ۴: مقایسه کارآمدی الگوریتم پیشنهادی با سایر روش ها

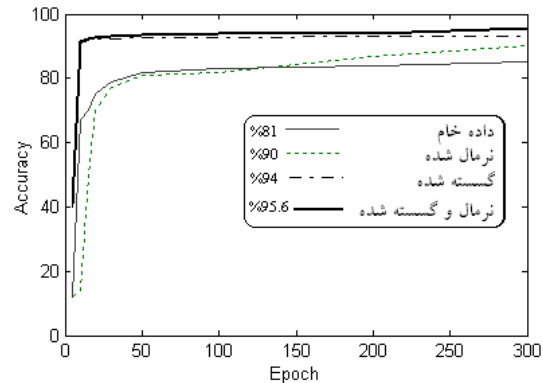
مرجع	نوع ویژگی	تعداد ویژگی	نرون لایه مخفی	دقت
۵	محلی مبتنی بر بلوک	۷۵	۲۰	٪۶۶
۶	سراسری	۳۰	۴۰	٪۸۱
الگوریتم پیشنهادی	محلی	۷۸	۲۰	٪۹۵.۶
	مبتنی بر ناحیه	۲۰	۲۰	٪۸۷

۷- نتیجه گیری

بازیابی تصویری مبتنی بر محتوا در دو مرحله انجام می پذیرد: دسته بندی و بازیابی. ما برای ارتقاء دقت سیستم های CBIR به افزایش دقت دسته بندی تصاویر پرداختیم. رویکرد ما برای افزایش دقت دسته بندی، استفاده از ویژگی هایی است که قابلیت دسته بندی دقیقتری داشته باشند. روش پیشنهادی یک الگوریتم سه مرحله ایست: «نرمال سازی، گسسته سازی و گزینش ویژگی». با نرمال سازی مقادیر بردار ویژگی از عدم آموزش ندیدن شبکه عصبی جلوگیری می شود. گسسته سازی، با وجود کاهش حجم داده ها از قدرت دسته بندی آنان نمی کاهد و در آخرین مرحله، ویژگی هایی که قدرت دسته بندی بیشتری دارند، گزینش شده و برای آموزش به شبکه عصبی ارائه می شوند. نتایج آزمایشی نشان می دهند الگوریتم پیشنهادی قادر است دسته بندی دقیقتری از تصاویر ارائه دهد.

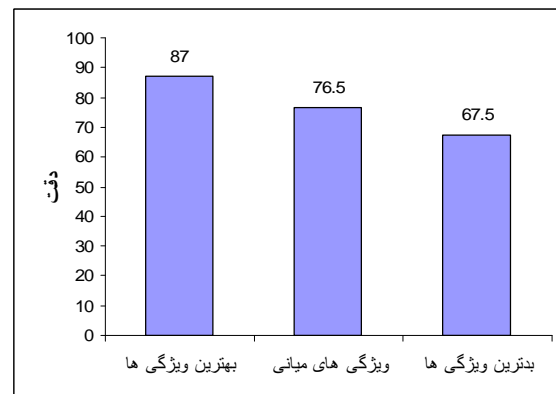
مراجع

- [1] J.Z. Wang, J. Li and G.Wiederhold, "SIMPLicity: Semantics-Sensitive Integrated Matching for Picture Libraries," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 9, pp. 947-963, 2001.
- [2] Ying Liu, Dengsheng Zhang, Guojun Lu and Wei-Ying Ma, "A survey of content-based image retrieval with high-level semantics", Journal of Pattern Recognition, 40(2007), 262- 282.
- [3] Wei Wang and Aidong Zhang, "Extracting semantic concepts from images: a decisive feature pattern mining approach", Multimedia Systems, 11(4), 352-366, 2006.
- [4] Yu Wang, Mingyue Ding, Chengping Zhou and Tianxu Zhang, "A Hybrid Method for Relevance Feedback in Image Retrieval Using Rough Sets and Neural Networks", international journal of computational cognition, vol. 3, no. 1, pp. 78-87, 2005.
- [5] Jun-Hua Han, D. S. Huang, Tat-Ming Lok, and Michael R. Lyu, "A novel image retrieval system based on BP neural network", IJCNN2005, Vol.4, no 31, pp.2561-2564, 2005.
- [6] H. young Ko and Suk In YOO: "Intelligent Image Retrieval Using Neural Network", IEICE TRANS. INF. & SYST, vol. E48-D, no.12, pp. 1810-1819, 2001.
- [7] B. C. Russell, A.Torralla, K. P.Murphy and W. T.Freeman, "LabelMe: a database and web-based tool for image annotation", MIT AI Lab Memo AIM-2005-025, 2005. www.labelme.csail.mit.edu
- [8] Witten, I.H. and Frank E., *Data Mining: Practical Machine Learning Tools and Techniques with JAVA Implementations*, Morgan Kaufman, San Francisco, 2005.
- [9] Tom. Mitchell, *Machine Learning*, McGraw-Hill, 1997.



شکل ۴: بررسی روند آموزش شبکه عصبی با چهار نوع داده ورودی: خام، نرمال شده، گسسته شده، نرمال و گسسته شده (در این آزمایش از گزینش ویژگی استفاده نشده است)

سپس ۲۰ عدد از بهترین ویژگی ها (دارای بیشترین بهره اطلاعات)، ۲۰ عدد از بدترین ویژگی ها (دارای کمترین بهره اطلاعات) و همچنین ۲۰ عدد از ویژگی های میانی را برگزیده و با هر یک، بطور جداگانه، به دسته بندی پرداختیم. نتایج این آزمایش را می توانید در شکل ۵ مشاهده فرمایید. نتایج آزمایشات فرضیه اولیه را تایید می کنند: «با گزینش تعدادی از بهترین ویژگی ها می توان به دسته بندی دقیقتری دست یافت».



شکل ۵: کیفیت ویژگی ها و تاثیر آنها در دقت دسته بندی

برای بررسی کارآمدی روش پیشنهادی، نتایج بدست آمده را با چند روش دیگر مقایسه نمودیم. تفاوت اصلی روش ما با روش ارائه شده در [5] در دو بخش است: ویژگی های مبتنی بر ناحیه و انتخاب ویژگی. چنانچه ملاحظه می شود این دو بخش در افزایش دقت دسته بندی تاثیر چشمگیری دارند. همچنین تفاوت اصلی روش پیشنهادی با [6]، در این است که ما به ارتقاء دقت دسته بندی و آنان به ارتقاء دقت بازیابی پرداختند. دقت نهایی ۸۱٪ نتیجه روش آنان می باشد