



شناسایی ارقام دستنویس فارسی با استفاده از الگوریتم خفاش

فاطمه صلاحی^۱

^۱کارشناسی ارشد، گروه کامپیوتر، دانشگاه آزاد اسلامی، بیرجند، ایران

Salahi.fateme@gmail.com

چکیده

شناسایی اعداد دستنویس یکی از موضوعات پیچیده در حوزه پردازش تصویر می باشد. با توجه به عصر دیجیتال، سرعت و دقت در سیستم های هوشمند پردازش تصویر اهمیت ویژه ای یافته است. از سوی دیگر با توجه به اهمیت شناسایی اعداد دستنویس و کاربردهای مختلف آن، در این پژوهش الگوریتمی به منظور شناسایی اعداد دستنویس فارسی ارائه شده است. در این پژوهش دو روش قاببندی و تبدیل موجک به منظور استخراج ویژگی استفاده شده است. همچنین از روش کلاسه بندی KNN استفاده شده است. درصد صحت روش قاببندی پس از تست ابعاد مختلف قاب بالای ۸۷ درصد بدست آمده است. با انتخاب ویژگیهای سطح سوم تبدیل موجک، درصد صحت ۹۵ بدست آمده است. با استفاده از الگوریتم تکاملی خفاش سعی در کاهش ویژگیها و بهبود سیستم شده است.

واژه های کلیدی: شناسایی اعداد دستنویس، تبدیل موجک، قاببندی، KNN، الگوریتم خفاش

۱. مقدمه

شناسایی دست نوشته ها یکی از فعالترین حوزه های تحقیقاتی در زمینه بینایی ماشین و پردازش الگو میباشد که به دلیل ویژگیهای خاص آن و همچنین کاربردهای بسیار زیاد مورد توجه بسیار قرار گرفته است. با توجه به اینکه حجم زیادی از اسناد کاغذی موجود، توسط اسکنرها یا دوربینها به اسناد تصویر دیجیتالی تبدیل میشوند، ذخیره سازی، بازیابی و مدیریت کارآمد این آرشیوهای تصویری، در بسیاری از برنامه ها نظیر اتوماسیون اداری و کتابخانه های دیجیتالی اهمیت فراوانی دارند. علاوه بر این، شناسایی رقم دست نوشته به عنوان یک معیار برای مقایسه طبقه بندیهای مختلف استفاده می شود. در تمامی این کاربردها مساله دقت و سرعت سیستم از درجه اهمیت بسیار بالایی برخوردار بوده و منجر به انتخاب این زمینه به عنوان پایه ای برای آزمون الگوریتم های پردازش الگو و دسته بندی اطلاعات شده است. استفاده از یک سیستم بازشناسی ارقام دست نویس در عمل، با چالش هایی مواجه است که مهمترین آنها ضرورت بالا بودن نرخ بازشناسی است.

امروزه، تشخیص دقیق کاراکترهای چاپی ماشینی تا حد زیادی یک مشکل حل شده است، با این حال، تشخیص دست نوشته ها نسبتاً دشوار است، زیرا افراد مختلف سبک دستخط متفاوت دارند. بنابراین، OCR دست خط هنوز یک موضوع تحقیقاتی فعال است. تشخیص کاراکترهای دست خط یک حوزه پژوهش در سه دهه گذشته بوده است (رضایی و کیمیایی، ۱۳۹۶). در نتیجه دستیابی به الگوریتم های موثر به منظور آنالیز تصویری اسناد، یک نیاز اساسی به حساب می آید. OCR به عنوان یکی از شاخه های فعال در شناسایی الگوی، آنالیز تصاویر اسناد و پردازش تصویر می باشد که تصاویر حاوی متن را به متن قابل جستجو تبدیل می کند.

روشهای بسیاری برای بازشناسی ارقام و حروف دستنویس لاتین پیشنهاد شده اند که از کارایی بسیار بالایی برخوردار هستند و میتوان گفت روشهای بازشناسی خودکار ارقام و حروف دستنویس لاتین به بلوغ خوبی رسیده است. در مقابل تحقیقات برای بازشناسی ارقام و حروف دستنویس فارسی/عربی پیشرفت چشمگیری نداشته است و تعداد پژوهشها در این حوزه بسیار اندک است (نادرزاد و شریفزاده، ۱۳۸۷). زبان فارسی، زبان گفتاری بیش از ۱۱۰ میلیون نفر از مردم، به طور عمده در ایران، افغانستان، تاجیکستان، و تا حدی در برخی کشورهای دیگر می باشد. همچنین زبانهای دیگری وجود دارند که از حروف و اعداد مشابه با زیرمجموعه آنها مانند عربی، اردو، پشتو استفاده میکنند، لذا انجام پژوهش در این حوزه لازم و ضروری است. در حوزه زبان فارسی، به دلیل شباهت زیادی که ارقام به هم دارند و همچنین تفاوت در شیوه رسم آنها، ایجاد یک سیستم بازشناسی با دقت قابل قبول برای استفاده عملی با مشکلاتی مواجه است. به همین دلیل، توسعه روشهایی برای بهبود دقت در آنها ضروری است.

شناسایی اعداد دستنویس دارای دو گام اساسی است. در گام اول سعی میشود تا اعداد از تصویر استخراج شده و از یکدیگر تفکیک گردند و در قدم بعدی مورد شناسایی قرار میگیرند. در خصوص قدم دوم تاکنون کارهای فراوانی صورت گرفته با این حال قدم اول به عنوان پیش درآمدی بر قدم دوم تلقی میشود. یکی از موثرترین کارهایی که به بهبود قابل توجه دقت در یک سیستم بازشناسی ارقام دستنویس می انجامد، استفاده از ویژگیهای مناسب برای بازنمایی ارقام است. این کار مستلزم انتخاب یک روش مناسب برای استخراج ویژگی و تعیین پارامترهای آن به صورت بهینه است. در این پژوهش برآنیم تا به مقایسه دو شیوه استخراج ویژگی قابیندی و تبدیل موجک و استفاده از الگوریتم خفاش^۱ به منظور یافتن بهترین ویژگیها بپردازیم.

این پژوهش در قالب ۵ بخش گردآوری شده است. در بخش بعدی به پیشینه تحقیقات صورت گرفته در این زمینه پرداخته شده است. در بخش سوم، روش تحقیق ارائه شده است. در بخش چهارم نتایج حاصل از پیاده سازی سیستم مطرح شده است. در بخش آخر خلاصه ای از نتایج بدست آمده مورد ارزیابی قرار گرفته و پیشنهاداتی برای ادامه یا بهبود سیستم ارائه شده است

۲. پیشینه پژوهش

در تشخیص دستنویس چه در زبان فارسی و چه انگلیسی پژوهش هایی انجام شده است. در پژوهشی (Surya et al, 2018) به شناسایی دست نوشته های انگلیسی با استفاده از شبکه عصبی عمیق^۲ پرداخته اند. در این پژوهش از دو پایگاه داده محبوب رقم و حرف انگلیسی، MNIST و EMNIST، برای تهیه مجموعه داده برای مرحله آموزش و آزمایش DNN استفاده شده است. دقت تشخیص برای ارقام و حروف انگلیسی به ترتیب 97/7٪ و 88/8٪ است. مقایسه عملکرد با سایر ساختارهای شبکه های عصبی نشان داد که میانگین وزنی تشخیص برای patternnet، feedforwardnet و DNN پیشنهادی به ترتیب 80/3٪، 68/3٪ و 9/4٪ بود.

در پژوهشی (Tohidul et al., 2018) به شناسایی ارقام دست نویس با استفاده از شبکه عصبی^۲ پرداخته اند. آزمایشها در این پژوهش بر روی پایگاه داده دست نویس MNIST که به صورت عمومی در دسترس است انجام شده است. شبکه عصبی مصنوعی چند لایه با عملکرد آزمون دارای دقت ۹۹/۶۰٪ است.

^۱ Bat Algorithm

^۲ Neural network

(Siddique et al., 2019) به شناسایی ارقام دست نویس با استفاده از شبکه عصبی کانولوشن^۳ پرداخته اند. هدف این مقاله مشاهده تنوع دقت CNN^۴ برای طبقه بندی ارقام دست نویس با استفاده از تعداد مختلف لایه ها و دوره های پنهان و مقایسه بین دقت ها است. برای این ارزیابی عملکرد CNN، الگوریتم پیشنهادی آنان بر مجموعه داده موسسه استاندارد و فناوری ملی اصلاح شده (MNIST) انجام شده است. شبکه با استفاده از گرادینان تصادفی نزول^۵ و الگوریتم تولید عقب^۶ آموزش داده می شود.

چن و همکاران (Chen et al., 2018) در پژوهش خود به شناسایی ارقام دستنویس انگلیسی با استفاده از آموزش عمیق^۷ پرداخته اند. در این مقاله عملکرد پنج مدل طبقه بندی یادگیری ماشین یعنی شبکه عصبی، k، نزدیکترین همسایه^۸ (K-NN)، جنگل تصادفی^۹، درخت تصمیم^{۱۰} و دسته بندی با افزایش شیب^{۱۱} مقایسه می شود. نتیجه این مقاله نشان می دهد که K-NN دارای دقت برابر ۹۶/۷٪ و شبکه عصبی دارای دقت ۹۶/۸٪ است، اما K-NN تقریباً ۱۰ برابر سریعتر به سرعت پردازش می رسد. این مقاله نشان می دهد که K-NN همراه با روش های پیش پردازش هنگامی که به عنوان یک الگوریتم طبقه بندی در شناسایی رقمی دست نویس آفلاین استفاده می شود، قادر به دستیابی به عملکرد عالی جدا از شبکه عصبی است.

جین و چوهان (Jain and Chauhan, 2018) در پژوهش خود به شناسایی ارقام دستنویس انگلیسی با استفاده از روشهای شبکه عصبی پرداخته اند. روش ارائه شده بر مبنای شبکه های عصبی عمیق و الگوریتم های یادگیری عمیق بوده است. الگوریتمهای NN^{۱۲} مانند، CNN.DNN^{۱۳}، و RNN^{۱۴} برای طبقه بندی ارقام دستنویس پیاده سازی شده اند. براساس نتایج بدست آمده، بهترین دقت مربوط به مدل CNN با درصد ۹۹/۶٪ بوده است.

در پژوهشی (ابوحمره و جاودانی، ۱۳۹۶) به بازشناسی مبلغ چک بانکی از روی مبلغ دستنویسته به عدد با استفاده از ویژگیهای آماری و ماشین بردار پشتیبان پرداخته اند. بر روی هر تصویر ۳ مرحله موجک هار^{۱۵} گرفته که در نهایت بردار ویژگی آن از تصویر تقریب سطح ۳ بدست می آید. برای طبقه بندی از ماشین بردار پشتیبان چند کلاسه به روش یکی در مقابل بقیه استفاده شده است. نتایج نشان داده که گاوسی درصد خطای کمتری دارد و درصد تشخیص ۹۶/۲ می باشد.

در پژوهش دیگر (محمدی، ۱۳۹۹) یک الگوریتم یادگیری با نظارت ضعیف برای شناسایی شماره پلاک خودرو با استفاده از شبکه های همگشتی عمیق^{۱۷} پیشنهاد داده اند. در فاز آموزش، تنها کاراکترهایی که در تصاویر پلاک وجود دارند مشخص میشود و الگوریتم پیشنهادی قادر است علاوه بر شناسایی وجود هر کاراکتر، مختصات آنها را نیز آشکارسازی نماید. به منظور ارزیابی عملکرد الگوریتم پیشنهادی، یک پایگاه داده شامل ۱۳۹۷ تصویر پلاک ایرانی جمع آوری شده است. نتایج آزمایشها نشان میدهد ۹۵/۶٪ از پلاکها و ۹۹/۱٪ از کاراکترها به درستی شناسایی شده اند.

^۳ convolutional neural network

^۵ stochastic gradient descent

^۶ backpropagation algorithm

^۷ Deep learning

^۸ K-nearest neighbor

^۹ Random Forest,

^{۱۰} Decision Tree

^{۱۱} Bagging with gradient boost

^{۱۲} Neural network

^{۱۳} Deep neural networks

^{۱۴} Convolutional neural networks

^{۱۵} Recurrent Neural network

^{۱۶} Haar

^{۱۷} Deep convolutional network

در تحقیقی (رضائی و کیمیایی، ۱۳۹۶) از یک سیستم منطق فازی برای تشخیص ارقام فارسی استفاده کرده اند. این روش مزایایی نسبت به نسخه غیرفازی اش دارا میباشد از جمله اینکه میزان اطمینان از دسته بندی یک نمونه را نشان میدهد، در مورد نمونه هایی که نسبت به آنها مردد است نظر نمیدهد و دارای درصد دسته بندی درست بالایی است.

در پژوهشی (فرمانی و سجودی، ۱۳۹۶) به بازشناسی ارقام و حروف دستنویس فارسی پرداخته اند. در مرحله استخراج ویژگیها، برای بهره مندی از مزایای مدل های عمیق از یک ساختار سلسله مراتبی دولایه ای استفاده شده است که بلوک سازنده آن، کدگذار خودکار نويز زدا میباشد. تابع کدکننده، داده های ورودی را به یک بازنمایش متفاوت تبدیل میکند و تابع دیکدکننده، بازنمایش جدید را به فرمت اولیه برمیگرداند. در مرحله طبقه بندی کننده از تابع تبدیل SoftMax استفاده شده است، این تابع تبدیل، برداری از المانهایی با مقدار حقیقی را دریافت و به برداری که هر یک از المانهایش در بازه (۰ و ۱) قرار میگیرد تبدیل میکند. براساس نتیجه بدست آمده دقت طبقه بندی براساس ویژگیهای پیشنهادی ۹۸/۳٪ گزارش شده است.

در پژوهشی دیگر (منتظر و همکاران ۱۳۹۵) از تبدیل ویژگیهای مقاوم در برابر تغییر مقیاس (سیفت¹⁸) و روش بسته ویژگیهای تصویری برای بازشناسی ارقام و حروف دستنویس فارسی/عربی استفاده می کنند. ابتدا ویژگیهای سیفت تصویر استخراج میشود. سپس به این ویژگیهای تصویری¹⁹، ویژگیهای هیستوگرام گردایانههای جهتدار²⁰ و ویژگیهای گابور²¹ نیز اضافه شده اند. از ویژگیهای استخراج شده برای آموزش دسته بندی کننده SVM استفاده میشود. روش پیشنهادی بر روی پایگاه داده های ارقام دستنویس HODA با نرخ تشخیص ۹۹/۶ درصد و پایگاه داده حروف دستنویس HODA با نرخ تشخیص ۹۴ درصد از روشهای دیگر پیشی گرفته است.

در تحقیقی (باسایی و حاتم، ۱۳۹۵) به انتخاب ویژگیهای موثر در بازشناسی ارقام دستنویس فارسی با الگوریتم تکاملی ژنتیک پرداخته اند. ویژگی ها از طریق گردایان در هشت جهت از هر تصویر استخراج می شود. برای طبقه بندی ارقام دست نویس از شبکه عصبی پرسپترون²² چند لایه استفاده شده است. استفاده از برخی عملیات پیش پردازش مانند نازک سازی²³، حذف نویزهای پس زمینه و پیش زمینه، استاندارد سازی کردن اندازه تصویر قبل از استخراج ویژگی می تواند در نتایج خروجی مفید باشد. صحت نتایج در استفاده از این روش بر روی پایگاه داده ی HODA برابر ۹۸/۸۵ درصد گزارش شده است.

پژوهش دیگری (فخری و رضوی، ۱۳۹۴) به تشخیص ارقام دستنویس فارسی با استفاده از یادگیری عمیق و ماشین بولتزمن²⁴ پرداخته اند. برای انجام و آزمایش نیاز به بانک اطلاعاتی ارقام دست نویس است که برای زبان انگلیسی بانک اطلاعاتی mnist و برای زبان فارسی بانک اطلاعاتی HODA موجود است. از ماشین بولتزمن عمیق و دولایه رمزگذار خودکار که در هر لایه از ۲۰۰ عدد نورون استفاده شده است. این روش نشان دهنده درصد بازشناسی اعداد تا ۹۳ درصد است.

در اقدامی دیگر (فرزاد، ۱۳۹۱) دو روش برای بهبود شناسایی ارقام دستنویس فارسی ارائه داده است. در مرحله ی استخراج ویژگی یک روش ترکیبی به نام هرم فضایی هیستوگرام گردایان²⁵ جهتدار و در مرحله ی طبقه بندی، یک ماشین بردار پشتیبان با دو هسته ی جدید برای بهبود دقت و زمان آموزش و آزمایش ارائه شده. آزمایشها بر روی پایگاه داده ی ارقام دستنویس HODA انجام شده و نتیجه ی آزمایشات، دقت ۹۹/۲۷ درصد را نشان میدهد که بهترین نرخ شناسایی کسب شده بر روی این مجموعه داده ها تاکنون میباشد.

¹⁸ SIFT (Scale Invariant Feature Transform)

¹⁹ BoVF (Bag of Visual Features)

²⁰ HOG (Histogram of Oriented Gradients)

²¹ GF (Gabor Features)

²² perceptron

²³ Thining

²⁴ Boltzmann Machine

²⁵ Pyramid Histogram of Oriented Gradient

پژوهشی (قنبری و همکاران، ۱۳۹۰) به ارائه یک روش انتخاب ویژگی مبتنی بر الگوریتم جستجوی گرانشی باینری²⁶ در سیستم بازشناسی ارقام دست نویس پرداخته اند. در این پژوهش یک روش هوشمند انتخاب ویژگی برای بازشناسی ارقام دست نویس فارسی ارائه شده است. در این روش تابع برازندگی مرتبط با خطای سیستم بازشناسی ارقام دست نویس فارسی با استفاده از الگوریتم جستجوی گرانشی باینری و با انتخاب ویژگی های مناسب کمینه می شود.

در تحقیقی (سلطانزاده و رحمتی، ۱۳۸۲) به شناسایی ارقام دستنویس فارسی با استفاده از گرادیان و دسته بندی کننده ماشین بردار پشتیبان پرداخته اند. ابتدا تصویر به اندازه استاندارد نرمال شده و گرادیان تصویر محاسبه میگردد. سپس برای هر نقطه از تصویر، زاویه گرادیان محاسبه شده و به ۴ یا ۸ زاویه استاندارد، تبدیل میگردد. از روی تصویر گرادیان حاصل، ۴ یا ۸ تصویر مجزا ساخته میشود که هر کدام از این تصاویر مقادیر گرادیان مربوط به یکی از زوایای استاندارد را در خود نگه میدارد. یا نمونه برداری از تصاویر ویژگیهایی نهایی استخراج شده اند. در این پژوهش عمل دسته بندی با استفاده از ماشینهای بردار پشتیبان صورت گرفته است. در این روش به میزان تشخیص ۹۹/۵۹ دست یافته اند.

۳. روش تحقیق

سامانه تشخیص کاراکتر نوری در تقسیم بندی دقیق به بخشهای پیش پردازش، پردازش، استخراج ویژگی، کاهش ویژگی، طبقه بندی و ارزیابی تقسیم می شود. در بعضی از سیستمهای تشخیص ارقام بخشهای استخراج ویژگی و طبقه بندی در هم ادغام میشوند. پیش پردازش کلیه اعمالی است که روی تصویر صورت میگیرند تا موجب تسهیل در روند اجرای فازهای بعدی گردد، مانند دوگانی کردن²⁷ تصویر، حذف نویز، هموارسازی، نازکسازی، تشخیص زبان و فونت کلمات، و نظایر اینها میباشد. مرحله پردازش در یک سیستم پردازش تصویر شامل چند مرحله می شود که عبارت است از استخراج ویژگی و کاهش ویژگیها که در ادامه به بررسی آنها پرداخته شده است.

استخراج ویژگی از مراحل بسیار با اهمیت در سیستمهای OCR است، چرا که نتایج حاصل از این مرحله، مستقیماً بر روی کیفیت مرحله بازشناسی اثر میگذارد. در انتخاب بردارهای ویژگی لازم است موارد چند مورد توجه قرار گیرند. بردار ویژگی هر الگو باید تا حد امکان از بردارهای ویژگی دیگر الگوها متمایز باشد. فاصله بین بردارهای ویژگی در فضای ویژگیهای حداکثر باشد. و بردار ویژگی الگوها باید تا بیشترین حد ممکن، خصوصیات شکل و ساختار الگوها را از تصویر آنها استخراج نماید. همچنین باید تا حد امکان نسبت به نویز، تغییر اندازه و نوع فونت، چرخش و دیگر تغییرات احتمالی الگوها دارای ثبات باشد. در مورد حروفی که به چندین شکل نوشته میشوند لازم است که بیش از یک کلاس الگو به یک کاراکتر خاص تعلق یابد.

۳-۱. روشهای استفاده شده برای استخراج ویژگی برای ارقام دستنویس

۳-۱-۱. تبدیل موجک

تبدیل موجک²⁸، یکی از روشهای انتخاب ویژگی در حوزه مشترک زمان و فرکانس است. در این روش، به جای استفاده از توابع سینوسی و کسینوسی (مانند تبدیل فوریه)، از توابعی به نام موجک استفاده میشود. هنگام استفاده از تبدیل موجک، تابع موجک شکل خود را حفظ میکند اما در طول سیگنال، جابه جا شده و نیز در طی جابه جا شدن، فشرده و باز میشود و بدین

²⁶ binary gravitational search algorithm.

²⁷ Binarization

²⁸ Discrete Wavelet Transform

ترتیب، کل سیگنال را در مینوردد. این روش برخلاف تبدیل فوریه زمان کوتاه، قادر است تا قدرت تفکیک متفاوتی را برای محدوده های فرکانس پایین و بالا ایجاد نماید.

رابطه تبدیل موجک پیوسته را می توان به صورت ضرب داخلی سیگنال و یک تابع پایه به فرم زیر نوشت. از ضرب داخلی $x(t)$ در نسخه ی انتقال یافته و مقیاس شده ی تابع $\Psi(t)$ که موجک نامیده میشود، تبدیل پیوسته ی موجک به دست میآید. از آنجا که در فرایند پایین تابع $\Psi(t)$ مقیاس می شود، اغلب آنالیز موجک را به جای آنالیز زمان-فرکانس، آنالیز زمان-مقیاس می نامند، هرچند آنالیز زمان-فرکانس به پارامتر مقیاس S بستگی دارد.

$$CWT_x^\Psi(\tau, s) = \Psi_x^\Psi(\tau, s) = \frac{1}{\sqrt{|s|}} \int_{-\infty}^{+\infty} x(t) \Psi^* \left(\frac{t-\tau}{s} \right) dt = \langle x(t), \Psi_{\tau, s}(t) \rangle \quad (1)$$

که در آن :

$$\Psi_{\tau, s}(t) = \frac{1}{\sqrt{|s|}} \Psi \left(\frac{t-\tau}{s} \right) \quad (2)$$

که در آن پارامترهای τ, s نشان دهنده ی کشیدگی یا مقیاس و انتقال (موقعیت زمانی یا مکانی) موجک نمونه هستند. ضرایب تبدیل موجک بیانگر میزان نزدیکی سیگنال به موجک در مقیاس مورد نظر است. در هر فضا بیش از یک مجموعه از توابع پایه وجود دارد که از بین آنها، توابع پایه متعامد از اهمیت ویژه ای برخوردارند چرا که خواص بسیار خوب و تسهیل کننده ای به ویژه در یافتن ضرایب تبدیل خواهند داشت. بدین سان، با استفاده از خاصیت تعامد توابع پایه، ضرایب تبدیل به صورت زیر محاسبه می شوند:

$$\alpha_k = \langle f(t), \phi_k(t) \rangle = \int f(t) \phi_k^*(t) dt \quad (3)$$

با داشتن این ضرایب، می توان تابع را به صورت زیر بازسازی نمود:

$$f(t) = \sum_{k=1}^N \alpha_k \phi_k(t) = \sum_{k=1}^N \langle f(t), \phi_k(t) \rangle \phi_k(t) \quad (4)$$

در کنار این خواص تسهیل کننده، ممکن است بسته به کاربرد، توابع پایه متعامد در دسترس نباشد. در این مواقع می توان از پایه های دومتعامد²⁹ استفاده نمود. در حالت کلی اگر پایه های دومتعامد نیز موجود نباشد، می توان از حالت کلی تری تحت عنوان فریم استفاده کرد.

تبدیل موجک، حاصل شباهت سنجی³⁰ بین محتوای فرکانسی (مقیاسی) سیگنال و تابع موجک در مقیاس های مختلف است. برای محاسبه تبدیل موجک پیوسته نیز پنجره مورد نظر منقبض/منبسط شده و شیفیت می یابد و در هر موقعیت، از حاصل ضرب آن در سیگنال، انتگرال زمانی گرفته می شود. در حالت گسسته، فیلترهایی با فرکانس قطع های مختلف برای تحلیل سیگنال در مقیاس های متفاوت به کار برده می شود. در حالت گسسته، رزولوشن سیگنال توسط عملکردهای فیلترها کنترل می شود و مقیاس از طریق Downsampling یا Upsampling تغییر می کند. روند پردازش با تبدیل ویولت گسسته چنین آغاز می شود؛ در ابتدا سیگنال از یک فیلتر دیجیتال پائین گذر نیم باند با پاسخ ضربه $h[n]$ عبور می کند، و لذا خروجی فیلتر برابر است با کانولوشن ورودی و پاسخ ضربه فیلتر؛ در نتیجه این عمل فیلترینگ، تمام مؤلفه های فرکانسی که بیشتر از نصف بزرگترین فرکانس موجود در سیگنال باشند حذف می شوند. از آنجا که بیشترین فرکانس موجود در سیگنال خروجی فیلتر برابر است با $\frac{\pi}{2}$ رادیان، نیمی از نمونه ها قابل حذف اند. لذا با حذف یکی در میان نمونه ها، طول سیگنال نصف

²⁹ biorthogonal

³⁰ Correlation

خواهد شد بدون این که اطلاعاتی را از دست داده باشیم. روند مشابهی نیز با استفاده از یک فیلتر دیجیتال بالاگذر نیم باند با پاسخ ضربه $g[n]$ انجام می پذیرد. در نتیجه در خروجی اولین مرحله از اعمال تبدیل موجک، دو نسخه، یکی بالاگذر و دیگری پائین گذر، با طول کاهش یافته (نصف شده) از سیگنال اولیه به فرم زیر بدست می آیند:

$$y_{high}[k] = \sum_n x[n].g[2k-n] \quad (5)$$

$$y_{low}[k] = \sum_n x[n].h[2k-n] \quad (6)$$

که در آن $h[n]$ پاسخ ضربه فیلتر پایین گذر نیم باند و $g[n]$ پاسخ ضربه فیلتر بالاگذر نیم باند میباشد. با این عمل، رزولوشن زمانی نصف شده و در مقابل رزولوشن فرکانسی دو برابر می شود. این روند را می توان مجدداً بر روی نسخه پائین گذر شده اعمال نمود و در هر مرحله، با کاهش رزولوشن زمانی به میزان نصف مرحله قبل، رزولوشن فرکانسی را دو برابر نمود. ضرایب خروجی فیلتر پائین گذر، شکل اولیه سیگنال را دنبال می کنند، به همین دلیل به این ضرایب، تقریب³¹ گفته می شود. همچنین ضرایب خروجی فیلتر بالا گذر، جزئیات فرکانس بالای سیگنال را دربردارند، به همین دلیل به این ضرایب، جزئیات³² گفته می شود. با افزایش تعداد مراحل تبدیل، میزان جزئیات نیز کاهش می یابد.

به منظور اعمال تبدیل موجک دوبعدی به تصویر، ابتدا تبدیل موجک یک بعدی به سطرها اعمال می شود و سپس ستونها با نرخ ۲، کاهش نمونه برداری³³ میشوند تا فقط نمونه های واقع در محل های زوج باقی بمانند. در این حالت، مجدداً تبدیل موجک، یک بعدی بر ستونها اعمال می گردد و نهایتاً سطرها با نرخ ۲، کاهش نمونه برداری میشوند. بدین ترتیب، ۴ زیرباند مختلف به عنوان ضرایب تبدیل موجک تصویر بدست می آید. (فرمانی و سجودی، ۱۳۹۶).

مشابه با حالت یک بعدی، اولین زیرباند از ضرایب تبدیل موجک مربوط به ضرایب تقریب است که از لحاظ مقدار و شکل ظاهری، مشابه با تصویر اولیه است. جدای از زیر باند تقریب، ۳ زیر باند جزئیات خواهیم داشت که یکی از آن ها مربوط به جزئیات افقی موجود در تصویر، یکی از آنها مربوط به جزئیات عمودی موجود در تصویر و آخرین، زیرباند مربوط به سایر جزئیات موجود در تصویر است که گاهی به آن، جزئیات قطری نیز گفته می شود.

۳-۱-۲. روش قاببندی

در فرایند استخراج ویژگی با استفاده از روش قاببندی، یک تصویر باینری جزبندی شده و به مقدار ثابتی از تصاویر کوچکتر که قاب نامیده می شوند شکسته می شود. از هر جعبه با این روش دو ویژگی استخراج میشود: یک مقدار ثابت برای هر جعبه که فاصله آن تا مبدا جعبه (گوشه پایین سمت چپ) برابر میانگین فاصله هر یک از پیکسلهای سیاه تا مبدا است. و ویژگی دوم یک زوایه که اندازه آن برابر میانگین اندازه زاویه های هر یک از پیکسلهای سیاه نسبت به سطح افق است. (کمالی و همکاران، ۱۳۸۶) که مطابق فرمول زیر میباشد:

$$r_b = \frac{1}{N} \sum_{k=1}^N d_k^N \quad b = 1.2 \dots 100 \quad (7)$$

$$a_b = \frac{1}{N} \sum_{k=1}^N \theta_k^N \quad b = 1.2 \dots 100 \quad (8)$$

³¹ . Approximation

³² . Detail

³³ downsample

که N تعداد پیکسلهای مورد نظر و d اندازه فاصله ی هر کدام از آنها تا گوشه سمت چپ پایین هر زیر تصویر و θ زاویه هر کدام از پیکسلها تا گوشه سمت چپ پایین هر زیر تصویر میباشد. اگر در زیر تصویری پیکسل مورد نظر وجود نداشته باشد اندازه و زاویه آن را برابر صفر قرار میدهیم.

۳-۲. روش کاهش ویژگی با الگوریتم خفاش

الگوریتم بهینه سازی خفاش الهامی از خصوصیات ردیابی خفاشهای کوچک در جستجوی شکار می باشد. به طوری که خفاشهای کوچک میتوانند در تاریکی مطلق با انتشار صدا و دریافت آن به شکار طعمه های خود بپردازند.

مکان x_i^t و سرعت v_i^t برای هر خفاش مجازی t ام در تکرار t و همچنین فرکانس f_i به صورت زیر محاسبه میگردد (بخته، ۱۳۹۳):

$$f_i = f_{min} + (f_{max} - f_{min})\beta \quad (9)$$

$$v_i^t = v_i^{t-1} + (x_i^t - x^*) \quad (10)$$

$$x_i^t = x_i^{t-1} + v_i^t \quad (11)$$

که در آن: $\beta \in [0.1]$ یک بردار تصادفی با توزیع یکواخت میباشد و x^* بهترین مکان فعلی است که در هر تکرار پس از مقایسه با موقعیت خفاشهای مجازی انتخاب میشود. معمولاً فرکانس f را با $f_{min} = 0$ و $f_{max} = 100$ در نظر می گیرند. در هر تکرار، در جستجوی محلی یکی از جوابها به عنوان بهترین جوابها انتخاب شده و موقعیت جدید هر خفاش به طور محلی با گام تصادفی به صورت زیر به روز میشود.

$$x_{new} = x_{old} + \epsilon A^t \quad (12)$$

که در آن $\epsilon \in [-1.1]$ یک عدد تصادفی بوده و $A^t = \langle A_i^t \rangle$ میانگین بلندی صدای خفاشها در تکرار t میباشد. همچنین بلندی صدای A_i^t و نرخ پالس ارسالی r در هر تکرار به صورت زیر به روز میشود: ($\gamma \cdot \alpha$ مقدار ثابت می باشند).

$$r_i^{t+1} = r_i^0 [1 - \exp(-\gamma t)] A_i^{t+1} = \alpha A_i^t \quad (13)$$

۳-۳. دسته بندی ارقام دستنویس با طبقه بندی کننده نزدیکترین همسایگی

این مرحله از شناسایی، شامل روشهایی برای متناظر ساختن هر یک از الگوهای به دست آمده از مرحله استخراج ویژگی ها، با یکی از کلاسهای فضای الگوهای مورد بحث است که از طریق کمینه ساختن فاصله بردار ویژگیهای هر الگوی ورودی نسبت به یکی از بردارهای مرجع، انجام میگردد. بردارهای مرجع، بردارهایی هستند که قبلاً از نمونه های آموزشی اخذ شده اند.

KNN یکی از متداولترین روشها برای دسته بندی داده ها میباشد. الگوریتم K همسایه نزدیک در میان سندهای آموزشی پیش دسته بندی شده، براساس یک معیار شباهت پیدا نموده و دسته های این K همسایه نزدیک برای پیش بینی دسته سند آموزشی به وسیله امتیازدهی سندهای هر دسته منتخب، استفاده خواهد شد (شفیعی و نخعی، ۱۳۹۷). یک نمونه ناشناخته براساس شباهت بین نمونه های شناخته شده آموزش دیده یا برچسب دار بر مبنای محاسبه فاصله بین نمونه های ناشناس با

نمونه های برچسبدار، شناخته میشود. سپس K نزدیکترین نمونه ها به عنوان پایه برای طبقه بندی انتخاب میشوند و نمونه نامشخص به کلاسی اختصاص مییابد که بیشترین نمونه ها را در میان نزدیکترین نمونه ها دارد. به همین منظور، الگوریتم طبقه بندی کننده KNN بستگی دارد به تعداد K همسایه عدد صحیح و تغییر مقدار پارامتر K که ممکن است نتایج طبقه بندی را تغییر دهد. و همچنین وابسته به مجموعه داده های برچسب دار می باشد. بنابراین اضافه کردن یا حذف هرگونه نمونه بر تصمیم نهایی طبقه بندی کننده KNN، تاثیر میگذارد. معیار فاصله، در KNN از فاصله اقلیدسی معمولاً به عنوان معیاری برای اندازه گیری فاصله بین دو نمونه استفاده میشود. طبقه بندی کننده KNN، به صورت تحلیلی قابل ردیابی است و به سادگی پیاده سازی میشود، اما یکی از مشکلات اصلی الگوریتم KNN این است که به همه نمونه های آموزشی نیاز دارد که در زمان اجرا در حافظه باشند، به همین دلیل، طبقه بندی مبتنی بر حافظه نامیده میشود (فهمیحسن و همکاران، ۱۳۹۷)

۳-۴. پایگاه داده Online-TMU

مجموعه داده ی Online-TMU توسط بخش مهندسی برق دانشگاه تربیت مدرس تهیه شده است. که مجموعه ی داده ی آموزشی و آزمایشی از دستخطها است. یکی از چالش هایی که در استفاده از این پایگاه وجود دارد عدم تفکیک بخشهای مختلف یک حرف مانند بدنه و نقطه هاست (مرزائی و همکاران، ۱۳۹۶). از چالش های دیگر این پایگاه داده عدم یکسان بودن سایز تصاویر میباشد که در این پژوهش با استفاده از یکنواخت کردن اندازه تصویر ابعاد تمامی تصاویر یکسان شده است.

۴. یافته ها

در پژوهش از پایگاه داده online-TMU استفاده شده است، که شامل ۱۰ کلاس و هر کلاس شامل یک نوع رقم میباشد. پس از انجام پیش پردازش روی داده ها، برای استخراج ویژگیها از دو روش استفاده شده است. روش اول مبتنی بر تبدیل موجک، و روش دوم با استفاده از قاببندی به استخراج ویژگیها پرداخته شده است. با استفاده از الگوریتم خفاش، سعی در انتخاب بهترین ویژگیها شده است. در انتها پس از استخراج ویژگیها از تصاویر تمام ارقام، با استفاده از روش کلاسبندی KNN به بررسی درصد خطا پرداخته شده است. در این پژوهش برای هر عدد تعداد ۱۵۰ تصویر انتخاب شده اند که از این میان ۱۰۰ تصویر به عنوان داده های آموزش و مابقی (۵۰ تصویر) به عنوان داده تست در نظر گرفته شده اند.

۴-۱. نرمال کردن اندازه تصویر در پیش پردازش

به دلیل اینکه ابعاد تصاویر ارقام موجود در پایگاه داده یکسان نیست، نیاز به نرمالیزه کردن تصاویر داریم. برای نرمال کردن تصویر با عرض w و طول l به تصویری با ابعاد مثلاً 100×100 ، اگر $L > W$ بود، L_{new} به ۱۰۰ نرمال شده و $w_{new} = 100 \cdot l / w$ و اگر $w > L$ بود w_{new} به ۱۰۰ نرمال و $L_{new} = 100 \cdot w / l$ نرمال میشود. سپس این تصویر در مرکز یک تصویر 100×100 که تمام درایه های آن هم رنگ پس زمینه تصویر اصلی نرمال شده هستند، قرار میگیرد. در شکل ۱ تصویر رقم یک قبل و بعد از نرمالسازی نشان داده شده است.



شکل ۱: تصویر عدد یک قبل و بعد از نرمالسازی

۴-۲. استخراج ویژگی

در این پژوهش استخراج ویژگی به دو روش (تبدیل موجک، قابندی) انجام شده است.

۴-۲-۱. استخراج ویژگی با استفاده از تبدیل موجک

در این مرحله استخراج ویژگی با استفاده از تبدیل موجک هار³⁴ صورت گرفته است. جهت اعمال دو سطح تبدیل موجک باید، بر روی سطح فرکانس پایین (LL) مجدد موجک را اعمال کرد. در این پژوهش از زیر باند فرکانس پایین به عنوان ویژگیهای هر عدد استفاده شده است. پس از دو بار اعمال موجک هار بر روی تصاویر، ابعاد ماتریس زیر باند فرکانس پایین، 25×25 میشود. سپس ماتریس های حاصله را با زیر هم قرار دادن ستون ها به بردار تبدیل میکنیم. در حقیقت ماتریس با ابعاد 25×25 به برداری با ابعاد ۶۲۵ تبدیل می شود. در ادامه به بررسی اعمال تبدیل موجک و بردار ویژگی استفاده شده برای یکی از تصاویر پایگاه داده پرداخته شده است.

پس از اعمال اولین سطح موجک هار بر روی تصاویر، ابعاد تصویر اصلی که 100×100 بوده، به 50×50 تبدیل شده است. در شکل ۲ اعمال سه سطح موجک بر روی زیر باند فرکانس پایین نشان داده شده است. در این تصاویر ابعاد ماتریس نتیجه، در سطح دوم، 25×25 و در سطح سوم 13×13 میباشد.



ج

ب

الف

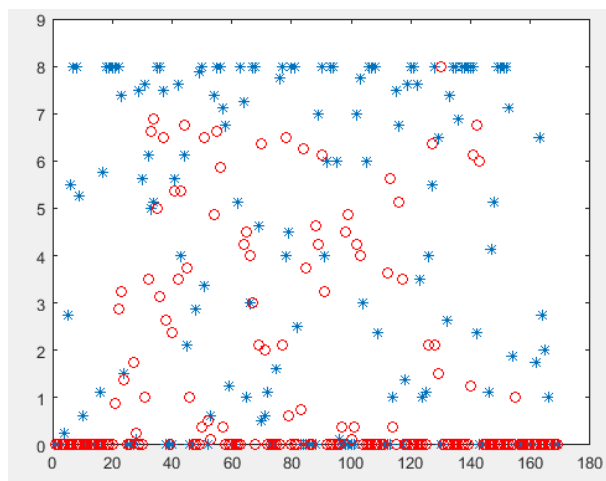
شکل ۲. اعمال تبدیل موجک به زیرباند فرکانس پایین. شکل الف. اعمال اولین سطح موجک هار به یکی از تصاویر پایگاه داده. از چپ به راست به ترتیب: LL، LH، HL، HH شکل ب نتیجه اعمال دومین سطح موجک هار، شکل ج نتیجه اعمال سومین سطح موجک هار.

به منظور بررسی ویژگیهای انتخاب شده در ادامه برای دو عدد صفر و پنج که شباهت زیادی در ظاهر به یکدیگر دارند، انواع روش های انتخاب ویژگی بر روی تصویر ۳ به کار برده شده است و نتایج هر یک مجزا نشان داده شده است. در ابتدا ویژگیها به کار برده شده بعد از سه سطح موجک هار نمایش داده شده است.



شکل ۳: دو تصویر از مجموعه پایگاه داده

³⁴ Harr



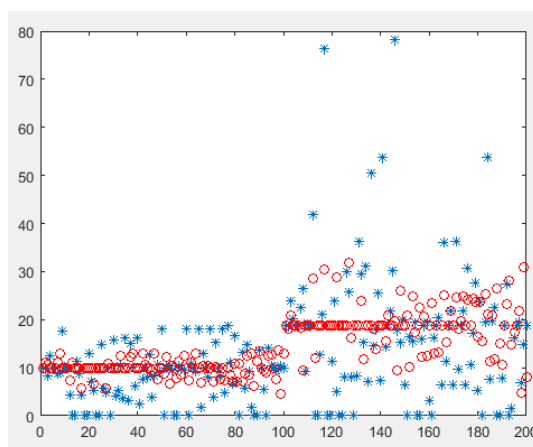
شکل ۴ ویژگیهای انتخاب شده توسط تبدیل موجک هار برای دو عدد(*) عدد صفر و ۵ برای عدد ۵)

باتوجه به اینکه سومین سطح موجک هار دارای ابعاد 13×13 میباشد، ابعاد ماتریس ویژگی برابر ۱۶۹ میباشد. همانطور که در شکل ۴ مشاهده می شود تعداد ۱۶۹ ویژگی برای هر عدد نمایش داده شده است. محور افقی شماره ویژگی ها و محور عمودی مقدار ویژگی می باشد. با توجه به شکل ۴، بردار ویژگی براساس تبدیل موجک تمایز خوبی میان اعداد ایجاد می کند.

۴-۲-۲. استخراج ویژگی با استفاده از قاببندی

در این مرحله استخراج ویژگی از تصاویر اعداد به روش قاببندی بوده است. در این الگوریتم، تصویر باینری هر عکس به ۱۰۰ زیرتصویر کوچکتر که قاب (زیر تصویر) نامیده میشود، تقسیم میگردد. سپس از هر کدام از این تصاویر ۲ ویژگی که شامل اندازه و زاویه می باشد استخراج خواهد گردید.

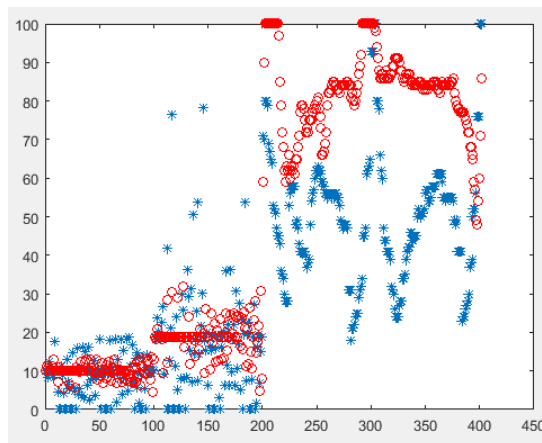
ابعاد بردار ویژگی حاصل شده در این مرحله، بستگی به انتخاب تعداد قابها در تصاویر هر عکس دارد. برای بدست آوردن کمترین میزان خطا در سیستم، آزمایشها با قابهایی با ابعاد مختلف انجام شده است که در بخشهای بعدی نتایج حاصل از این آزمایشها بطور کامل بررسی شده است و بعد قابهایی که بالاترین دقت را در سیستم دارند، مشخص شده است. به منظور بررسی قدرت تمیز ویژگی توسط این روش، در شکل زیر به مقایسه ویژگیها برای اعداد شکل ۳ پرداخته شده است.



شکل ۵: ویژگیهای انتخاب شده براساس روش قاببندی برای دو عدد(*) عدد صفر و ۵ برای عدد ۵)

در اینجا ابعاد هر قاب ۲۰×۵ انتخاب شده است. و با توجه به ابعاد اصلی ماتریس (۱۰۰×۱۰۰) تعداد قابها برابر با ۱۰۰ میباشد. از آنجا که برای هر قاب دو ویژگی استخراج میشود تعداد ویژگیها برابر با ۲۰۰ میباشد.

به منظور ایجاد تمایز بیشتر میان ویژگیها، چند ویژگی جدید به این روش اضافه شده است. برای هر تصویر، جمع هر سطر، جمع هر ستون، جمع قطر اصلی و جمع قطر فرعی تصویر به عنوان ویژگی جدید به ویژگیهای قبلی اضافه شده است. خروجی این مرحله یک بردار ویژگی ۴۰۲×۱ میباشد.



شکل ۶. ویژگیهای انتخاب شده براساس روش قاببندی و اضافه کردن ویژگیهای جدید برای دو عدد (* عدد صفر و O برای عدد ۵)

۳-۴. طبقه بندی براساس K-NN و نتایج

در این مرحله با استفاده از الگوریتم K-NN که در آن $K=۳$ انتخاب شده است (این مقدار براساس اجرای چندین بار الگوریتم و مقایسه نتایج برای K مختلف بهترین نتیجه را حاصل کرده است) به طبقه بندی و محاسبه درصد خطا پرداخته شده است. از هر عدد تعداد ۱۰۰ تصویر به عنوان آموزش و ۵۰ تصویر به عنوان تست، انتخاب شده اند. در الگوریتم KNN به منظور محاسبه فاصله میان بردارهای ویژگی از فاصله اقلیدسی استفاده شده است. بدین صورت که فاصله هر بردار تست با بردارهای آموزش محاسبه شده است و کمترین فاصله ها انتخاب شده اند.

۴-۴. نتایج انتخاب ویژگی بر مبنای تبدیل موجک

ابتدا با استفاده از بردار ویژگی بدست آمده از تبدیل موجک شبکه را آموزش داده ایم. همانطور که از مقایسه نتایج جدول ۱ مشخص است، بهترین نتیجه برای سطح سوم موجک ۲db با درصد صحت ۹۵٪ بدست آمده است.

جدول ۱. نتایج حاصل از اعمال سه موجک مختلف در سه سطح

موجک 2db		موجک 1sym		موجک هار		
ابعاد بردار ویژگی	درصد صحت	ابعاد بردار ویژگی	درصد صحت	ابعاد بردار ویژگی	درصد صحت	
2601	91/4%	2500	92%	2500	92%	سطح اول موجک
729	94/2%	625	94/2%	625	94/2%	سطح دوم موجک
225	95%	169	94/6%	169	94/6%	سطح سوم موجک

۴-۵. نتایج انتخاب ویژگی بر مبنای روش قاببندی

در این مرحله، براساس انتخاب ویژگی بر مبنای روش قاببندی مجدداً سیستم را آموزش داده و نتایج حاصل به شرح جدول زیر میباشد. ابتدا فقط ویژگیهای مبنی بر روش قاببندی مورد استفاده قرار گرفته اند. همان طور که در جدول ۲ نشان داده شده است، با استفاده از ابعاد مختلف برای قاببندی درصد صحت متفاوت بوده است و بهترین درصد صحت در این روش استفاده از ابعاد قاب 10×10 می باشد. در مرحله بعد، فقط بردارهای مربوط به جمع سطر، ستون و قطر اصلی و فرعی به عنوان ویژگیها مورد استفاده قرار گرفته اند. سپس هر دو این ویژگیها به عنوان بردار ویژگی به سیستم اعمال شده اند. همانطور که از مقایسه نتایج بدست آمده از سه روش فوق مشخص است، استفاده از هر دو نوع ویژگی در کنار یکدیگر بهترین نتیجه را حاصل میکند. همچنین بکارگیری ابعاد 10×10 برای قاببندی نتیجه بهتری در مقایسه با سایر ابعاد دارد. با این حال، نتیجه بدست آمده از انتخاب ویژگی بر مبنای تبدیل موجک در مقایسه با روش قاببندی بهتر بوده است. به همین سبب در ادامه از این روش استفاده شده است.

جدول ۲. نتایج بدست آمده از روش قاببندی

		ابعاد قاب	ابعاد بردار ویژگی	نتیجه (درصد صحت)
روش ۱	روش قاببندی	10*10	200	81%
		5*20	200	74/8%
		20*5	200	75/8%
روش ۲	ویژگی ها : جمع سطر و ستون و قطر ها		202	80%
روش ۳	روش قاببندی و سطر و ستون	10*10	402	87/2%
		5*20	402	86%
		20*5	402	84/6%

۴-۶- بکارگیری الگوریتم خفاش

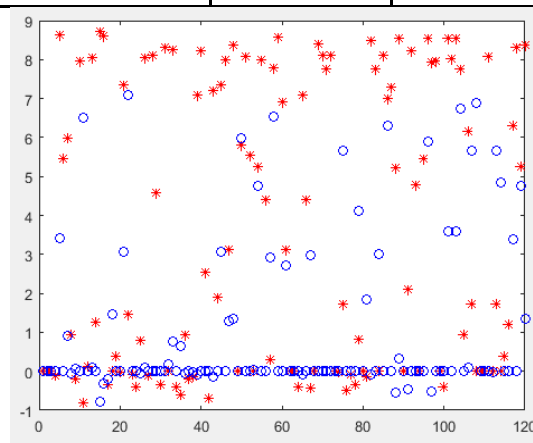
از آنجایی که انتخاب ویژگی نقش مهمی در پردازش تصاویر دارد، در این پژوهش از الگوریتم خفاش به منظور کاهش بردار ویژگی و انتخاب بهترین ویژگی ها استفاده شده است که باعث کاهش در حجم ذخیره سازی می شود و زمان آموزش و استفاده را کاهش می دهد؛ تا مدلی با پیچیدگی کمتر را به وجود بیاورد؛ که این خود موجب افزایش کارایی خواهد شد. تابع هدف در الگوریتم خفاش، ماکزیمم شدن میزان درصد صحت میباشد. بطوریکه در هر نسل، ویژگیها به سمتی در حرکت هستند تا میزان صحت افزایش یابد. در این الگوریتم تعداد نسل تکرار 90 انتخاب شده است. همچنین بلندی صدا برابر 0/5 در نظر گرفته شده است و نرخ پالس در ابتدا 0/5 قرار دارد. جمعیت اولیه 30 می باشد. پس از طی کردن 90 تکرار الگوریتم به جواب بهینه دست می یابد.

۴-۶-۱. بکارگیری الگوریتم خفاش در روش تبدیل موجک

در ابتدا الگوریتم خفاش بر روی ویژگی های استخراج شده از تبدیل موجک اعمال شد. در جدول به طور خلاصه به ویژگیهای الگوریتم خفاش پرداخته شده است.

جدول ۳: مشخصات الگوریتم خفاش به کار رفته برای ویژگی تبدیل موجک

تعداد نسل تکرار	تعداد ویژگیها	تعداد جمعیت اولیه
90	120	30



شکل ۷. ویژگیهای انتخاب شده توسط الگوریتم خفاش در روش استخراج ویژگی تبدیل موجک برای دو عدد

(* عدد صفر و O برای عدد ۵)

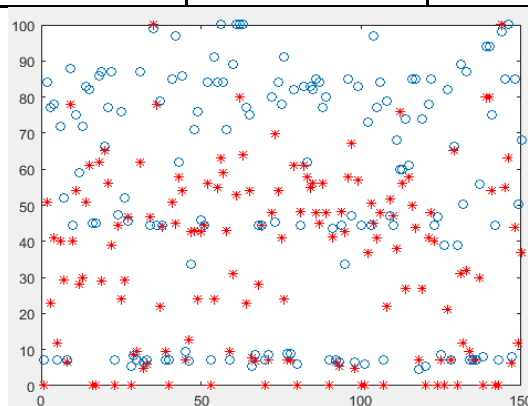
شکل ۷ نتیجه ی حاصل از انتخاب ویژگی ها توسط الگوریتم خفاش در روش تبدیل موجک بر روی دو عدد شکل ۳ می باشد. این الگوریتم پس از تکرار 90، توانسته ابعاد بردار ویژگی را کاهش دهد. به طوری که با استفاده از تبدیل موجک در سطح سوم در بهترین حالت آن به تعداد 225 عدد ویژگی رسیده بودیم، اما با استفاده از این الگوریتم ابعاد بردار ویژگی به میزان چشمگیری به عدد 120 کاهش پیدا کرده است. بطوریکه اگر از میان 225 ویژگی که از روش تبدیل موجک به درصد صحت 95 رسیده ایم، تعداد 120 ویژگی را انتخاب کنیم، سیستم همچنان کارایی بالای خود را حفظ خواهد کرد.

۴-۶-۲. بکارگیری الگوریتم خفاش در روش قاببندی

در ادامه الگوریتم خفاش برای انتخاب ویژگیهای استخراج شده از روش قاببندی مورد تست قرار گرفته است. همان طور که بیان شد استخراج ویژگی بر مبنای روش قاببندی توانست ابعاد بردار ویژگی را در بهترین حالت به ۴۰۲ عدد برساند. جدول زیر به طور خلاصه مشخصات الگوریتم خفاش به کار رفته را نشان می دهد.

جدول ۴: مشخصات الگوریتم خفاش به کار رفته برای ویژگی روش قاببندی

تعداد نسل تکرار	تعداد ویژگیها	تعداد جمعیت اولیه
90	150	30



شکل ۸: ویژگیهای انتخاب شده توسط الگوریتم خفاش در روش استخراج ویژگی قاببندی برای دو عدد (*) عدد صفر و 0 برای عدد ۵)

شکل ۸ ویژگی های انتخاب شده توسط این الگوریتم پس از اعمال آن در روش قاببندی را نشان می دهد. این شکل نتیجه ی پردازش بر روی دو عدد بیان شده در ابتدای فصل می باشد. همان طور که بیان شد در بهترین حالت قاببندی با اندازه قاب ۱۰*۱۰ به تعداد ۴۰۲ ویژگی رسیده بود اما این الگوریتم توانست تعداد این ویژگی ها را به میزان قابل توجهی کاهش دهد؛ طوری که پس از اعمال این الگوریتم در روش استخراج ویژگی با استفاده از قاببندی تعداد ویژگی ها به عدد ۱۵۰ رسید.

۵. نتایج

با توجه به نتایج بدست آمده در بهترین حالت در هر یک از روش ها، مطابق جدول ۵، چنین نتیجه گرفته میشود که روش استخراج ویژگی با استفاده از تبدیل موجک، در مقایسه با روش قاببندی عملکرد بهتری داشته است. همچنین الگوریتم خفاش به منظور کاهش ابعاد بردار ویژگی به کار گرفته شده است. کاهش ویژگیها در سیستم تشخیص سبب افزایش سرعت سیستم، کاهش پیچیدگی و بهبود عملکرد سیستم میگردد.

جدول ۵: مقایسه ی تعداد ویژگی در روش های مختلف

ابعاد بردار ویژگی در بهترین حالت قبل از اعمال الگوریتم خفاش	ابعاد بردار ویژگی پس از اعمال الگوریتم خفاش	روش استخراج ویژگی
۴۰۲	۱۵۰	قاببندی
۲۲۵	۱۲۰	تبدیل موجک

در این پژوهش برای شناسایی ارقام دستنویس فارسی پایگاه داده Online-TMU، از دو روش استخراج ویژگی (تبدیل موجک، و قاببندی) استفاده شده است. به منظور کلاسبندی از الگوریتم KNN استفاده شده است. براساس نتایج بدست آمده با استفاده از روش قاببندی ۸۷/۳ درصد صحت بدست آمده است، اما با روش تبدیل موجک با سه سطح موجک 2db به نتیجه درصد صحت ۹۵ رسیده ایم. این درصد صحت با تعداد ویژگی ۲۲۵ بدست آمده. سپس با استفاده از الگوریتم خفاش به کاهش ویژگیها پرداخته شده است. و به نتیجه کاهش تعداد ویژگیها به میزان ۱۰۰ دست یافته ایم. الگوریتم به کار گرفته شده پیچیدگی کمی داشته و سرعت بالایی دارد و در عین حال درصد صحت بدست آمده درصد بالایی بوده و قابلیت اطمینان خوبی دارد. با توجه به انجام مطالعات و شبیه سازیهای انجام شده پیشنهادات زیر برای ادامه کار داده میشود: استفاده از روشهای کاهش ویژگی مانند، PCA، استفاده از روشهای دیگر دسته بندی مانند شبکه عصبی و SVM، اضافه کردن شناسایی حروف فارسی دستنویس به سیستم، استفاده از روش یادگیری عمیق^{۳۵}، ارزیابی روش برای رزولوشن های مختلف تصاویر

³⁵ Deep learning

فهرست مراجع

۱. ابوحمزه، رامین و جاودانی، محسن (۱۳۹۶). بازشناسی مبلغ چک بانکی از روی مبلغ دست نوشته به عدد با استفاده از ویژگیهای آماری و ماشینبردار پشتیبان. هشتمین همایش سالانه بانکداری الکترونیک و نظامهای پرداخت. تهران مرکز همایشهای بین المللی برج میلاد.
۲. الهی، مژده (۱۳۹۲). شناسایی چهره با استفاده از تبدیل موجک و SVM. پایان نامه کارشناسی ارشد. دانشگاه آزاد گناباد.
۳. بخته، سمیه (۱۳۹۳). الگوریتم فراابتکای خفاش برای مسائل برنامه ریزی عدد صحیح. هشتمین کنفرانس تحقیق در عملیات.
۴. حسن زاده، م. (۱۳۹۱). بهبود نرخ بازشناسی ارقام دست نویس فارسی با استفاده از روش های ادغام در سطح تصمیم، پایان نامه کارشناسی ارشد رشته برق-الکترونیک، دانشگاه بیرجند.
۵. رضائی، عباسعلی، کیمیایی، فرزانه (۱۳۹۶). استفاده از منطق فازی برای تشخیص ارقام فارسی. مجله فناوری اطلاعات در طراحی مهندسی. دوره دهم. شماره دوم.
۶. سلطان زاده، حسن و محمد رحمتی، ۱۳۸۲، شناسایی ارقام دستنویس فارسی با استفاده از گرادیان و دسته بندی کننده ماشین بردار پشتیبان، نهمین کنفرانس سالانه انجمن کامپیوتر ایران، تهران، دانشگاه صنعتی شریف،
https://www.civilica.com/Paper-ACCSI09-ACCSI09_041.html
۷. شفیعی، شیماء؛ نخعی، ناهید (۱۳۹۷). تکنیکی با هدف بهبود سرعت و دقت الگوریتم دسته بندی کننده KNN. مجله نخبگان علوم و مهندسی. جلد ۳. شماره ۵.
۸. شیخ الاسلامی، مهدی؛ شیخالسلامی، مصطفی (۱۳۹۵). فشرده سازی تصویر با استفاده از تکنیکهای داده کاوی. چهارمین کنفرانس بینالمللی پژوهش در علوم و تکنولوژی. سنپترزبورگ-روسیه .
۹. صدیقیناو، محسن؛ سلیمانیاوری، علی؛ خسروی، حسین (۱۳۹۳). کاهش ویژگی توسط توده ذرات دودویی برای بازشناسی ارقام دستنویس فارسی. هوش محاسباتی در مهندسی برق. سال پنجم. شماره اول.
۱۰. علیپور، مهدی؛ نیکفرجام، کاظم (۱۳۹۶). مروری بر الگوریتمهای استخراج ویژگی و روشهای طبقه بندی در سیستمهای بازشناسی ارقام دستنویس فارسی برون خط. اولین کنفرانس ملی کاربردهای فناوریهای نوین در علوم و مهندسی برق، کامپیوتر و IT.
۱۱. فرخی، علیرضا و سیدناصر رضوی، ۱۳۹۴، تشخیص ارقام دست نویس فارسی با استفاده از یادگیری عمیق و ماشین بولتزمن، هفتمین کنفرانس ملی مهندسی برق و الکترونیک ایران، گناباد، دانشگاه آزاد اسلامی گناباد،
https://www.civilica.com/Paper-ICEEE07-ICEEE07_083.html---11
۱۲. فرزاد، فرهاد (۱۳۹۱). یک روش بهبود یافته برای استخراج ویژگی و شناسایی ارقام دستنویس فارسی. پایان نامه جهت اخذ کارشناسی ارشد. مهندسی کامپیوتر-معماری. دانشگاه شهید بهشتی. معاونت فناوری اطلاعات و ارتباطات. مدیریت آموزش الکترونیکی.
۱۳. فرمانی، آرزو و سجودی، امید (۱۳۹۶). بازشناسی ارقام و حروف دستنویس فارسی با استفاده از یادگیری عمیق بازنمایشها. دومین کنفرانس بینالمللی پژوهشهای دانشبنیان در مهندسی کامپیوتر و فناوری اطلاعات. تهران.

۱۴. فرهادیانی، رامین؛ صفری، عبدالرضا؛ همایونی، سعید (۱۳۹۸). کاهش اسپیکل در تصاویر راداری با دریچه ی گشایش مصنوعی درحوزهی موجک با بکارگیری وابستگیهای درون مقیاسی و بین مقیاسی. نشریه علمی-پژوهشی علوم و فنون نقشهبرداری. دوره هشتم. شماره ۴.
۱۵. فلاحپور، سعید؛ ریکنده، جوادعلیپور (۱۳۹۳). پیشبینی شاخص سهام با استفاده از شبکه های عصبی موجکی در بورس اوراق بهادار تهران. راهبرد مدیریت مالی. سال دوم. شماره ۷. دانشکده علوم اجتماعی و اقتصادی. دانشگاه الزهرا
۱۶. فهمیحسن؛ آرش؛ مغاری، محمدرضا؛ عبادتی، امیدمهدی (۱۳۹۷). پیشبینی اهدا خون با استفاده از داده کاوی بر پایه الگوریتمهای درخت تصمیم، SVM.KNN، و MLP. دوفصلنامه مدیریت مهندسی و رایانش نرم. سال چهارم، شماره ۱. ۷۷-۹۷
۱۷. قجر، ج. و سریانی، م. و کوچاری، ع. (۱۳۸۹). تشخیص چهره با استفاده از PCA و فیلتر گابور، دو فصل نامه پردازش علائم و داده ها، شماره ۱، صفحه ۸۹-۹۶.
۱۸. قنبری، نجمه؛ رضوی، سیدمحمد؛ نبوی کریزی، سیدحسن (۱۳۹۰) ارائه یک روش هوشمند انتخاب ویژگی مبتنی بر الگوریتم جستجوی گرانشی باینری در سیستم بازشناسی ارقام دست نویس فارسی. نشریه مهندسی برق و مهندسی کامپیوتر ایران. سال ۶. شماره ۱.
۱۹. کمالی، محمد؛ ایزدیان، جلالالدین؛ قوچانی، سعید؛ خیرخواه، احمدرضا (۱۳۸۶). بازشناسی حروف دستنویس به وسیله سیستم فازی به کمک روش قابندی در استخراج ویژگی. اولین کنگره مشترک سیستمهای فازی و سیستمهای هوشمند .
۲۰. محمدی، محمدرضا (۱۳۹۹). شناسایی شماره پلاک خودرو بر اساس یادگیری عمیق با نظارت ضعیف. ماشین بینایی و پردازش تصویر. مجله انجمن ماشین بینایی و پردازش تصویر ایران.
۲۱. مرزائی، مجید؛ رضوی، سیدمحمد؛ تقیپور، مهران (۱۳۹۶). بازشناسی برخط ارقام دستنویس فارسی با استفاده از طبقه بند ماشینبردار پشتیبان. مجله علمی پژوهشی رایانش نرم و فناوری اطلاعات. جلد ۶. شماره ۱.
۲۲. منتظر، غلامعلی؛ سلطاننهایی، محمدعلی؛ گیوکی، داور (۱۳۹۵). روش جدیدی برای بازشناسی حروف و ارقام دستنویس فارسی بر پایه تبدیل ویژگیهای مقاوم در برابر تغییر مقیاس. بیست و چهارمین کنفرانس مهندسی برق ایران. دانشگاه شیراز. اردیبهشت ۱۳۹۵
۲۳. نادرزاد، احسان؛ شریفزاده، سارا (۱۳۸۷). کاهش نویز اسپیکل تصاویر با استفاده از معادلات انتشار و مفهوم تصویر پیکسونی. چهاردهمین کنفرانس ملی سالانه انجمن کامپیوتر ایران. دانشگاه صنعتی امیرکبیر. ایران. تهران.
۲۴. یاسایی، ساناز و احمد حاتم، (۱۳۹۵)، انتخاب ویژگی های مؤثر در بازشناسی ارقام دستنویس فارسی با الگوریتم تکاملی ژنتیک، اولین کنفرانس بین المللی دستاوردهای نوین پژوهشی در مهندسی برق و کامپیوتر، تهران، کنفدراسیون بین المللی مخترعان جهان (IFIA)، دانشگاه جامع علمی کاربردی، -https://www.civilica.com/Paper-CBCONF01-CBCONF01_0200.html
25. Chen,S. Almamlook,R. Gu,Y. Well,L (2018). Offline Handwritten Digits Recognition Using Machine learning.Proceedings of the International Conference on Industrial Engineering and Operations Management Washington DC, USA, September 27-29, 2018
26. Jain,S. Chauhan,R (2018). Handwritten Digit Recognition using DNN, CNN and RNN. International Conference on Advances in Computing and Data Sciences.
27. Siddique,Fathma. Sakib,Shadman, Siddique,Abu Bakr (2019). Recognition of Handwritten Digit using Convolutional Neural Network in Python with Tensorflow and Comparison of Performance for

Various Hidden Layers. To be published in 5th International Conference on Advances in Electrical Engineering (ICAEE-2019)

28. Surya,T.Razi Mohd Noor,A. Hartiwi,M. (2018). Development of English Handwritten Recognition Using Deep Neural Network. Indonesian Journal of Electrical Engineering and Computer Science. Vol. 10, No. 2, May 2018, pp. 562~568. ISSN: 2502-4752, DOI: 10.11591/ijeecs.v10.i2.pp562-568

29. Tohidul.Kh, Gopal.Ram, Mujataba.Ghulam, Friday.Henry(2018). Handwritten Digits Recognition with Artificial Neural Network. Proc. of the International Conference on Engineering Technologies and Technopreneurship (ICE2T 2017) 18-20 September 2017, Kuala Lumpur, Malaysia

Identification of Persian handwritten cultivars using the bat algorithm

Fatemeh Salahi

Department of Electrical and Computer Engineering, Islamic Azad University, Birjand, Iran

Salahi.fateme@gmail.com

Abstract

Recognition of handwritten numbers is a complex issue in the field of image processing. Due to being in the digital age, speed and accuracy of intelligent systems have always been developing, which is why image processing has become particularly important. On the other hand, due to the importance of handwritten numbers recognitions and its various applications, in this study, an algorithm has been proposed to recognize Persian handwritten numbers. In this research, two methods of “framing and wavelet transform” have been used to extract the feature. The KNN classification method is also used. After testing different dimensions of the mold, the accuracy percentage of the framing method has been obtained above 87% By selecting the features of the third level of wavelet transform, 95% accuracy is obtained. Using the evolutionary bat algorithm, it has been tried to reduce the features and improve the system.

Keywords: handwritten numbers recognition, wavelet transform, framing, KNN, bat algorithm