



تشخیص بیماری نیوکاسل مرغ با استفاده از آنالیز موجک صدای قلب

محمد صادقی^۱، احمد بناکار^{۱*} و عبدالحمید شوشتری^۲

۱- به ترتیب دانشجوی کارشناسی ارشد و استاد یار گروه مهندسی مکانیک بیوسیستم دانشکده کشاورزی دانشگاه تربیت مدرس.

۲- استادیار مرکز تحقیقات واکسن و سرم سازی رازی کرج

ایمیل مکاتبه کننده: ah_banakar@modares.ac.ir

چکیده

در این پژوهش به تشخیص بیماری نیوکاسل پرنده از روی سیگنال صدای قلب پرداخته شده است. بدین منظور جوجه‌های مرغ ابتدا به دو دسته تقسیم شدند. یک گروه توسط بیماری نیوکاسل آلوده گردید و توسط گوشی پزشکی به ضبط صدای قلب جوجه‌ها پرداخته شد. این سیگنال‌ها توسط تبدیل موجک دابچی نوع اول به حوزه زمان-فرکانس انتقال داده شدند و ویژگی‌ها از ضرایب تقریب و جزئیات سطح دوم تبدیل موجک استخراج گردید. به منظور دسترسی به بهترین ویژگی‌ها با استفاده از روش IDE، تعداد ۱۲ ویژگی انتخاب شد. ماشین بردار پشتیبان (SVM) به عنوان طبقه بند مورد استفاده قرار گرفت و تابع شعاعی گوسی (RBF) به منظور ابر صفحه این طبقه بند به کار گرفته شد. در مرحله اول ماشین بردار پشتیبان با ۱۷۲ سیگنال صدای قلب آموزش داده شد تا ساختار مناسب طبقه بند شکل گیرد و پس از آن با ۴۲ سیگنال صدای قلب مورد آزمون قرار گرفت. دقت SVM در تشخیص بیماری نیوکاسل بر اساس سیگنال صدای قلب ۲ روز پس از اعمال بیماری ۹۲/۸۶٪ به دست آمد.

واژه‌های کلیدی: بیماری نیوکاسل، سیگنال صدای قلب مرغ، ماشین بردار پشتیبان، داده کاوی، تبدیل موجک

مقدمه

بیماری نیوکاسل یک بیماری مسری بین پرندگان است که هم در گونه‌های اهلی و هم در گونه‌های وحشی شایع می‌باشد. این بیماری قابلیت سرایت به انسان را نیز دارا است. این بیماری برای اولین بار در سال ۱۹۲۶ در نیوکاسل انگلستان شناسایی شد و از آن پس بدین نام مشهور گردید (Morgan, 1946). عامل ویروس بیماری نیوکاسل (NDV)، ویروس تک رشته RNA می‌باشد. NDV از اعضای خانواده پارامیکسویروس یک پاتوژن اصلی مرغ می‌باشد که با به خطر انداختن



سلامتی مرغ، مشکلات اقتصادی فراوانی را سبب می‌شود (Nidzworski et al., 2013). این بیماری از طریق مدفوع و سایر ترشحات پرندگان آلوده و همچنین از طریق تماس با مواد غذایی آلوده، آب، تجهیزات و لباس انتقال پیدا می‌کند. بستر آلوده تا ۲ ماه و لاشه آلوده تا ۱۲ ماه ویروس را نگهداری می‌کنند. عوامل مختلفی از جمله مدفوع بستر، ترشحات بزاق و مواد پروتئینی اطراف ویروس در افزایش مقاومت ویروس نقش اساسی دارند. کالبد گشایی بر اساس شدت ویروس، سن ابتلاء و نژاد صورت می‌گیرد. از علائم گوارشی می‌توان به خشکی لاشه، تیرگی لاشه، خونریزی‌های چربی قاعده قلب و از علائم تنفسی می‌توان به ترشحات سروزی در بینی و نای همراه با خونریزی اشاره نمود (Alexander, 1988 ; Capua et al., 2009). عوامل مختلفی از جمله بیماری، ضعف و گونه ممکن است در سیگنال صوتی منتشره از پرندگان موثر باشد. پژوهش‌های متفاوتی در زمینه سیستم‌های زیستی بر اساس روش‌های هوش مصنوعی و داده کاوی صورت پذیرفته است. آکودو و کورادا (۲۰۰۹) به منظور تشخیص و طبقه‌بندی خودکار ۳ گونه پرندگان و ۹ قورباغه از روی کمینه و بیشینه فرکانس، طول صدا و بیشینه توان، از ماشین بردار پشتیبان استفاده کردند. در این پژوهش به منظور بررسی بیشتر، از درخت تصمیم و آنالیز جداساز خطی نیز استفاده شد که ماشین بردار پشتیبان توانست با دقت ۹۵٪ به طبقه‌بندی بپردازد. در این راستا دقت درخت تصمیم و آنالیز جداساز خطی به ترتیب ۸۹٪ و ۷۱٪ بدست آمد (Acevedo et al., 2009). ابراهیم تورکوگلو و همکاران در سال ۲۰۰۳ یک سیستم هوشمند بر اساس تشخیص الگو ارائه کردند که سیگنال داپلر از بیماری‌های دریچه قلب انسان را تفسیر می‌کند. مدل ارائه شده متشکل از دو لایه موجک و پرسپترون چند لایه بود. لایه موجک به منظور استخراج ویژگی تطبیقی در حوزه زمان-فرکانس و پرسپترون چند لایه هم به منظور طبقه‌بند مورد استفاده قرار گرفت. دقت طبقه‌بندی شبکه عصبی موجک ۹۴٪ بدست آمد (Turkoglu et al., 2003). علیزاده و همکاران در سال ۱۳۹۲ با استفاده از شبکه عصبی مصنوعی و با دقت ۹۳٪ به تشخیص امراض دریچه‌ای قلب انسان بر اساس سیگنال PCG پرداختند. آن‌ها بر اساس تبدیل موجک دابچی مرحله چهارم و استفاده از ضرایب تقریب و جزئیات به استخراج ویژگی‌های سیگنال مبادرت ورزیدند و تشخیص بیماری را به وسیله شبکه عصبی پرسپترون دو لایه انجام دادند (علیزاده و همکاران، ۱۳۹۲).

در این پژوهش روشی مبتنی بر هوش مصنوعی و پردازش سیگنال به منظور تشخیص بیماری نیوکاسل مرغ از روی صدای قلب مورد بررسی قرار گرفته است. بدین منظور از ماشین بردار پشتیبان (SVM) به عنوان طبقه‌بند استفاده گردید. برای دسترسی به بیشینه دقت طبقه‌بند، از روش کاهش ابعاد و انتخاب ویژگی IDE استفاده شد. با استفاده از روش IDE، ویژگی‌هایی که مناسب‌ترین شاخص برای شناسایی یک سیگنال هستند، انتخاب گردیدند. سیگنال صدای قلب در حوزه زمان-فرکانس با استفاده از تبدیل موجک گسسته (DWT) دابچی نوع اول تا سطح دوم مورد پردازش واقع شد. در ادامه به تشریح روند اجرای مراحل مختلف پژوهش و نتایج به دست آمده پرداخته خواهد شد.

۲- مواد و روش‌ها

۲-۱ تبدیل موجک

سیگنال‌های برداشته شده توسط حسگرهای صدا ابتدا در حوزه زمان هستند. سیگنال‌های حوزه زمان دارای اهمیت ویژه‌ای می‌باشند. در واقع هر گونه آنالیز ثانویه و داده کاوی بر اساس سیگنال‌های این حوزه صورت می‌پذیرد. اما از سیگنال‌های



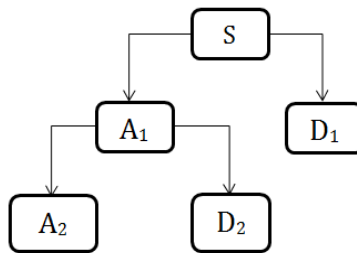
حوزه زمان نمی‌توان به ماهیت فرکانسی سیگنال پی برد بنابراین به منظور دستیابی به ماهیت دوره‌ای یک سیگنال به پردازش آن در حوزه فرکانس نیاز است (Lee et al., 1999). علاوه بر آن در برخی موارد محتوای فرکانسی یک سیگنال اطلاعات بهتری را به منظور شناسایی آن سیگنال در اختیار قرار می‌دهد (Duhamel et al., 1990). روش تبدیل فوریه سریع در پردازش سیگنال‌های ناپایدار که فرکانس آنها در طول زمان تغییر می‌کند خوب عمل نمی‌کند (Wang et al., 2009) و این عیب ناشی از آن است که در تبدیل فوریه سریع شفافیت زمانی صفر است (Iyer et al., 2012). از طرفی تحلیل‌های فرکانسی و زمانی هیچ کدام به تنهایی قادر به ارائه هم‌زمان اطلاعات فرکانسی و زمانی از یک سیگنال نیستند. سیگنال صدای قلب یک نوع از شکل موج اسیلاتوری در حوزه زمان است، در حالی که فرکانس آن با زمان تغییر می‌کند (نخعی و همکاران، ۱۳۹۲). بنابراین لازم است که به هر دوی اطلاعات حاصله از سیگنال در حوزه زمان و فرکانس توجه نمود. تبدیل موجک گسسته یک تحلیل دو بعدی سیگنال به منظور رسیدن هم‌زمان به شفافیت زمانی و فرکانسی می‌باشد و روشی بسیار موثر در پردازش سیگنال محسوب می‌گردد. در تبدیل موجک با تمامی مولفه‌های فرکانسی یک سیگنال مانند تبدیل فوریه سریع یکسان رفتار نمی‌شود بلکه هدف اصلی تبدیل موجک ارائه شفافیت زمانی دقیق و فرکانسی نادقیق برای تغییرات تند و سریع؛ و شفافیت فرکانسی دقیق و زمانی نادقیق برای تغییرات کند و آهسته است (Misiti et al., 1996). تبدیل موجک در آنالیز سیگنال‌هایی که دارای نویز می‌باشند با توجه به ماهیت حذف ذاتی موج دارای مزیت می‌باشد (Marchant, 2003). به منظور پردازش سیگنال‌ها در حوزه زمان-فرکانس از دو نوع تبدیل موجک پیوسته و گسسته استفاده می‌شود. از آنجا که یک سیگنال پیوسته دارای بی نهایت مقدار است که وارد کردن آن‌ها به فرمول تبدیل موجک دشوار و نیازمند محاسبات پیچیده می‌باشد، برای نیل به قابلیت بارگذاری و کدگذاری تبدیل موجک بر روی رایانه از تبدیل موجک گسسته (DWT) استفاده می‌شود. DWT توسط فرمول زیر تعریف می‌گردد (Saravanan et al., 2010):

$$DWT \tilde{\psi}(j, k) = \Psi \tilde{\psi}(j, k) = \frac{1}{\sqrt{2^j}} \int_{-\infty}^{+\infty} x(t) \psi\left(\frac{t-2^j k}{2^j}\right) dt \quad (1)$$

کمیت ψ در رابطه ۱، تابعی است که دارای طول معین، میانگین صفر، رفتار نوسانی، کاملاً میرا و وظیفه پنجره بندی سیگنال را بر عهده دارد. در واقع کمیت ψ که موجک مادر نام دارد، الگویی است جهت بازسازی سیگنال اولیه. با استفاده از تبدیل موجک گسسته، یک سیگنال به دو مولفه فرکانس پایین و فرکانس بالا تقسیم بندی می‌شود. در واقع رابطه ۱ به صورت دو فیلتر بالاگذر و پایین گذر، سیگنال را به تقریب‌ها و جزئیات تجزیه می‌نماید. در تبدیل موجک گسسته، تقریب معرف شکل اصلی سیگنال است و جزئیات بیانگر محدوده‌های فرکانس بالای نهفته در سیگنال اصلی هستند (Misiti et al., 1996). در این پژوهش به منظور پردازش سیگنال صدای قلب جوجه سالم و بیمار از DWT استفاده شد. منابع مختلفی کارا بودن DWT را در کاربردهای مختلف از جمله تشخیص بیماری‌های قلبی بیان نموده‌اند (علیزاده و همکاران، ۱۳۹۲؛ نخعی و همکاران، ۱۳۹۲). به منظور استخراج ویژگی از سیگنال‌های قلبی، خانواده دابچی در مقایسه با سایر موجک‌ها بسیار مناسب‌تر است و از قدرت بیشتری در تحلیل سیگنال‌های گذرا و غیر ایستا برخوردار می‌باشد (علیزاده و همکاران، ۱۳۹۲).



در این پژوهش با استفاده از گوشی پزشکی Littman 3M که قابلیت ضبط صدای قلب را داراست به ضبط صدای قلب جوجه مرغ پرداخته شد و سیگنال‌های صدا با استفاده از موجک مادر دابچی نوع اول و تا ۲ مرحله پردازش گردیدند. شکل ۱ معرف یک تبدیل موجک گسسته تا ۲ مرحله است.



شکل ۱- تبدیل موجک گسسته پیاده سازی شده در ۲ مرحله

۲-۲ داده کاوی

۲-۲-۱ استخراج ویژگی

ضرایب تبدیل موجک (یا خروجی هر نوع پردازنده سیگنال) نمی‌توانند به صورت مستقیم به عنوان ورودی طبقه‌بند مورد استفاده قرار بگیرند. چرا که این مقادیر گروه بسیار بزرگی از داده‌های خام را در بر می‌گیرند و تعبیر ریاضی-فیزیکی خاصی نیز ندارند. توابع ویژگی توابعی هستند که هر کدام به نوبه خود حالتی از یک سیگنال را بیان می‌کنند و حاوی اطلاعاتی هستند که به عنوان ورودی طبقه‌بند از آن‌ها می‌توان بهره جست. این ویژگی‌ها بر روی سیگنال‌های پردازش شده در هر سه حوزه زمان، فرکانس (تبدیل فوریه سریع) و زمان-فرکانس (تبدیل موجک گسسته) تعریف می‌شوند و به منظور بررسی دقیق‌تر رفتار سیگنال‌های صدا مورد استفاده قرار می‌گیرند. ۲۳ ویژگی مورد استفاده در این پژوهش در جدول ۱ نشان داده شده است (Lei et al., 2008; Khazaei et al., 2013). در این جدول $x(n)$ معرف سیگنال به ازای $n=1,2,\dots,N$ است که N تعداد نقاط داده‌ها می‌باشد.



جدول ۱- ویژگی‌های استخراج شده از سیگنال صدای قلب جوجه‌ها

فرمول ویژگی‌ها		
$T_1 = \frac{\sum_{n=1}^N x(n)}{N}$	$T_9 = \frac{T_5}{T_1}$	$T_{17} = \frac{\sum_{n=1}^N (x(n) - T_1)^4}{(N-1)T_1^4}$
$T_2 = \max x(n) $	$T_{10} = \frac{T_2}{\frac{1}{N} \sum_{n=1}^N x(n) }$	$T_{18} = \frac{1}{N} \sum_{i=1}^N (x_i - T_2)^3$
$T_3 = \sqrt{\frac{\sum_{n=1}^N (x(n) - T_1)^2}{N-1}}$	$T_{11} = \frac{\sum_{n=1}^N (x(n) - T_1)^2}{N-1}$	$T_{19} = \frac{1}{N} \sum_{i=1}^N (x_i - T_2)^4$
$T_4 = \left(\frac{\sum_{n=1}^N \sqrt{ x(n) }}{N} \right)^2$	$T_{12} = \frac{N}{\sum_{i=1}^N \frac{1}{x_i}}$	$T_{20} = \frac{1}{N} \sum_{i=1}^N (x_i - T_2)^5$
$T_5 = \sqrt{\frac{\sum_{n=1}^N (x(n))^2}{N}}$	$T_{13} = \sqrt[3]{\prod_{i=1}^N x_i}$	$T_{21} = \frac{1}{N} \sum_{i=1}^N (x_i - T_2)^6$
$T_6 = \frac{T_{18}}{T_3^2}$	$T_{14} = \frac{T_3}{T_2} \times 100$	$T_{22} = \frac{T_{19}}{T_{17}^2}$
$T_7 = \frac{T_2}{T_3}$	$T_{15} = \sum_{i=1}^N \frac{(x_i - T_2)}{N}$	$T_{23} = \sum_{n=1}^N (x(n))^2$
$T_8 = \frac{T_2}{\frac{1}{N} \sum_{n=1}^N x(n) }$	$T_{16} = \frac{T_{18}}{T_3^2}$	

۲-۲-۲ انتخاب ویژگی

انتخاب ویژگی بایستی از دو جنبه کیفی و کمی بودن به شدت مورد توجه باشد. انتخاب تعداد فراوان توابع ویژگی سبب سردرگم شدن طبقه‌بند می‌گردد و طبقه‌بند نمی‌تواند بین دو گروه ویژگی استخراج شده از دو طبقه سیگنال، تمایزی قائل شود. کیفیت ویژگی از آن جنبه حائز اهمیت است که ویژگی انتخاب شده بایستی برای شناسایی سیگنال مفید باشد (Bagheri et al., 2010). در این پژوهش از روش IDE به منظور انتخاب بهترین ویژگی‌ها استفاده شده است. پارامترهای به کار رفته در این روش عبارتند از:

$$\{q_{m,c,j}, m=1,2,\dots,M_c; c=1,2,\dots,C; j=1,2,\dots,J\} \quad (3)$$

که مقدار ویژه زمین ویژگی از m امین نمونه مربوط به کلاس c می‌باشد. M_c تعداد نمونه‌های c امین کلاس و J تعداد ویژگی‌های هر کلاس است. با در نظر گرفتن پارامترهای فوق روابط IDE به شکل زیر تعریف می‌گردند (Lei et al., 2008):



مرحله ۱: محاسبه فاصله متوسط نمونه‌های درون کلاس مشخص:

$$d_{c,j} = \frac{1}{M_c \times (M_c - 1)} \sum_{l,m=1}^{M_c} |q_{m,c,j} - q_{l,c,j}|$$

$$l, m = 1, 2, \dots, M_c, \quad l \neq m; \quad (4)$$

سپس بدست آوردن فاصله متوسط بین مقادیر یک ویژگی در هر کلاس:

$$d_j^{(w)} = \frac{1}{C} \sum_{c=1}^C d_{c,j} \quad (5)$$

هر چقدر فاصله متوسط بین ویژگی‌های یک کلاس کمتر باشد، موید این مطلب است که ویژگی‌ها برای آن کلاس از پراکندگی کمتری برخوردار هستند؛ به عبارت دیگر هر چقدر این پارامتر برای یک ویژگی کوچک‌تر باشد، نشان دهنده مناسب‌تر بودن آن ویژگی برای شناسایی سیگنال مربوط به آن کلاس است.

مرحله ۲: تعریف واریانس مربوط به فاکتور $d_j^{(w)}$ مطابق فرمول زیر:

$$v_j^{(w)} = \frac{\max(d_{c,j})}{\min(d_{c,j})} \quad (6)$$

مرحله ۳: این مرحله شامل محاسبه فاصله متوسط مقادیر ویژگی‌های بین کلاس‌ها است. قبل از آن میانگین مقادیر ویژگی‌های هر کلاس به صورت زیر تعریف می‌شود:

$$u_{c,j} = \frac{1}{M_c} \sum_{m=1}^{M_c} q_{m,c,j} \quad (7)$$

سپس فاصله متوسط بین مقادیر ویژگی‌های بین کلاسی محاسبه می‌گردد:

$$d_j^{(b)} = \frac{1}{C \times (C - 1)} \sum_{l,m=1}^{M_c} |u_{e,j} - u_{c,j}|$$

$$c, e = 1, 2, \dots, C_c, \quad c \neq e; \quad (8)$$

در واقع در این مرحله فاصله مقادیر ویژگی‌های یک کلاس از مقادیر ویژگی‌های کلاس دیگر بدست می‌آید. هر چقدر این فاصله بیشتر باشد، شاخص مناسب‌تری برای تمایز بین دو کلاس است.

مرحله ۴: تعریف واریانس مربوط به فاکتور $d_j^{(b)}$ مطابق فرمول زیر:

$$v_j^{(b)} = \frac{\max(|u_{e,j} - u_{c,j}|)}{\min(|u_{e,j} - u_{c,j}|)} \quad (9)$$

مرحله ۵: تعریف و محاسبه فاکتور پاداش:



$$\lambda_j = \frac{1}{\frac{v_j^{(w)}}{\max(v_j^{(w)})} + \frac{v_j^{(b)}}{\max(v_j^{(b)})}} \quad (10)$$

مرحله ۶: محاسبه امتیاز هر ویژگی بر اساس نسبت بین $d_j^{(w)}$ و $d_j^{(b)}$ با در نظر گرفتن فاکتور پاداش مطابق رابطه زیر:

$$\alpha_j = \lambda_j \times \frac{d_j^{(b)}}{d_j^{(w)}} \quad (11)$$

از فرمول ۱۱ می‌توان برداشت نمود که ویژگی‌ای بیشترین امتیاز را دارد که بیشترین اختلاف برون کلاسی و کمترین اختلاف درون کلاسی را داشته باشد.

مرحله ۷: در نهایت با استفاده از فرمول زیر، امتیاز هر ویژگی نرمال شده و با در نظر گرفتن حد آستانه دلخواه بهترین ویژگی‌ها انتخاب می‌شوند:

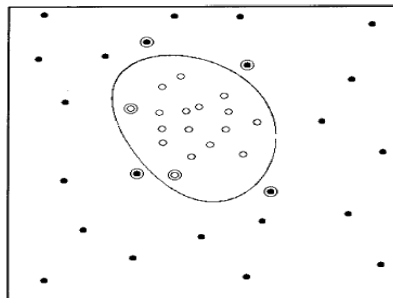
$$\bar{\alpha}_j = \frac{\alpha_j}{\max(\alpha_j)} \quad (12)$$

۲-۳ ماشین بردار پشتیبان

ماشین بردار پشتیبان یک روش بسیار قدرتمند در طبقه‌بندی است که اولین بار توسط کورتس و وپنیک در سال ۱۹۹۵ میلادی بر پایه تئوری یادگیری آماری بنا نهاده شد (Cortes et al., 1995). ایده اصلی در این روش طبقه‌بندی جدا نمودن کلاس‌ها با استفاده از یک ابر صفحه فرضی است. شکل ۲ مفهوم ابر صفحه را به وضوح بیان می‌دارد. در این طبقه بند ماکزیمم کردن حاشیه بین دو کلاس مد نظر می‌باشد. در این پژوهش از تابع پایه شعاعی گوسی (RBF) به عنوان جداکننده و ابر صفحه استفاده شد. RBF با فرمول زیر تعریف می‌شود (Burgess, 1998):

$$K(x, y) = \exp\left(-\|x - y\|^2 / 2\sigma^2\right) \quad (13)$$

که در آن x و y به ترتیب به فضای ویژگی و فضای هدف اشاره دارند و σ در مفهوم پهناى ابر صفحه به کار می‌رود.



شکل ۲- مفهوم ابر صفحه در ماشین بردار پشتیبان (Scholkopf et al., 1997)



جوجه‌های ۱۴ روزه با تلقیح ویروس نیوکاسل به صورت چشمی آلوده گردیدند. با استفاده از گوشی پزشکی Littman 3M (مدل ۳۲۰۰ با قابلیت فیلتر کردن در سه بازه فرکانسی $20-200\text{Hz}$ ، $500-1000\text{Hz}$ و $50-500\text{Hz}$ و قطر دیافراگم ۵/۱ سانتی‌متر) در روزهای ۱۵ و ۱۶ صدای قلبشان ضبط گردید. شکل ۳ مراحل صورت گرفته در این پژوهش به منظور تشخیص بیماری نیوکاسل پرنده از روی صدای قلب را نشان می‌دهد. در ادامه به بیان نتایج در روز ۱۶ (دو روز بعد از اعمال بیماری) پرداخته می‌شود.

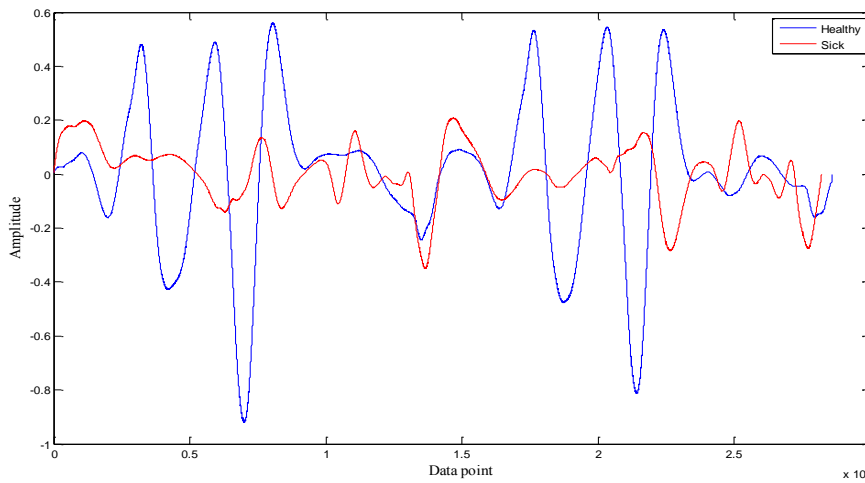


شکل ۳ - فرآیند تشخیص بیماری نیوکاسل

۳-نتایج و بحث

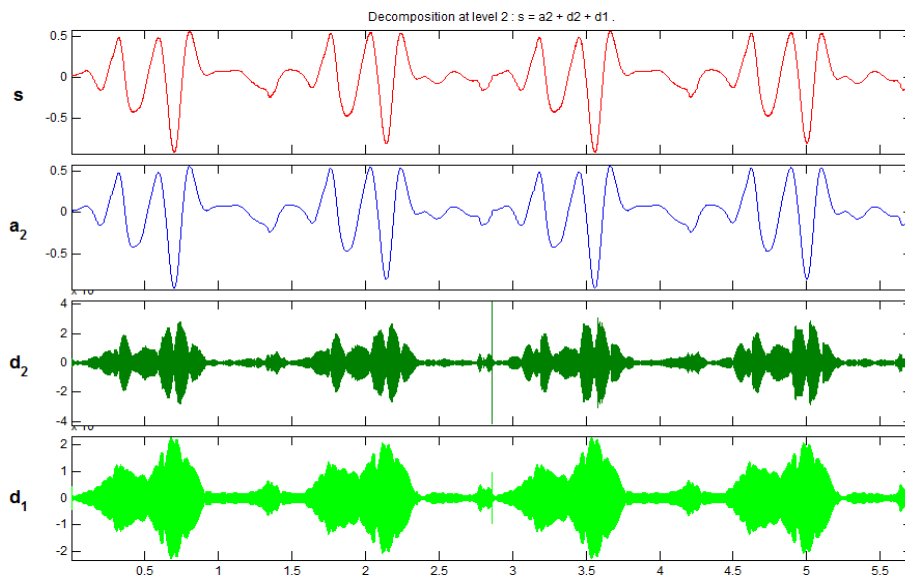
۳-۱ سیگنال‌های تحصیل شده

در این مرحله به نمایش سیگنال‌های بدست آمده در حوزه زمان-فرکانس پرداخته می‌شود. شکل ۴ معرف یک نمونه سیگنال صدای قلب جوجه بیمار و سالم در حوزه زمان است.

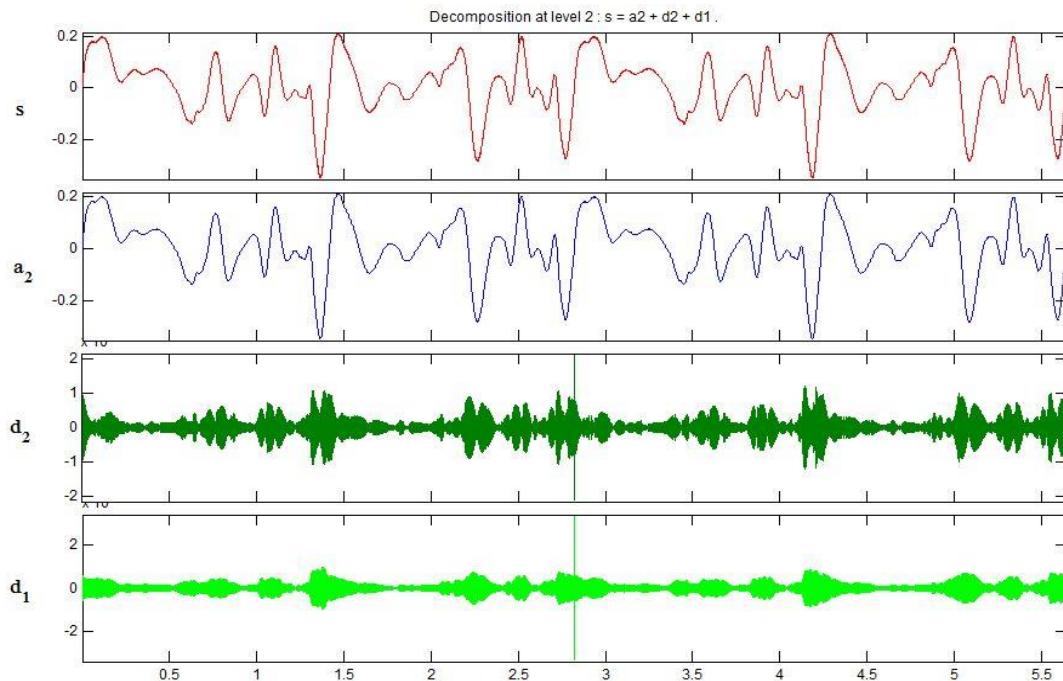


شکل ۴ - نمونه سیگنال صدای قلب جوجه سالم و بیمار، Healthy جوجه سالم. Sick جوجه بیمار.

سیگنال‌های صدای قلب توسط تبدیل موجک دایچی نوع اول در سطح دوم به حوزه زمان-فرکانس برده شدند. شکل‌های ۵ و ۶ به ترتیب معرف سیگنال‌های مربوط به حوزه زمان-فرکانس صدای قلب جوجه سالم و بیمار هستند. از شکل‌های ۴، ۵ و ۶ به وضوح مشخص است که صدای قلب جوجه سالم از شدت بیشتری برخوردار است. همچنین از ضرایب جزئیات شکل‌های ۵ و ۶ آشکار است که صدای قلب جوجه سالم یکنواخت‌تر از صدای قلب جوجه بیمار است.



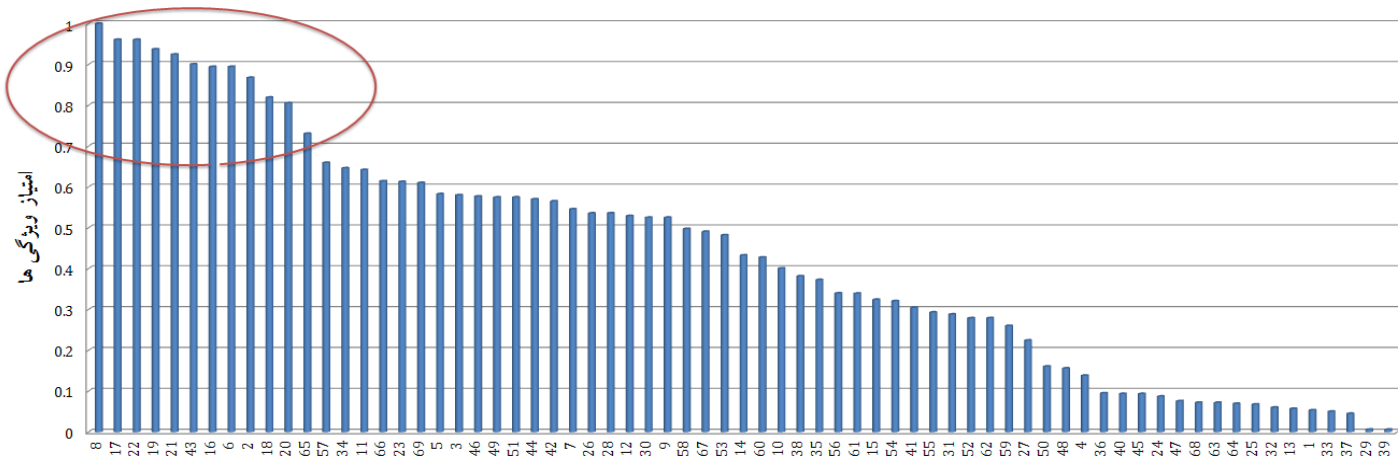
شکل ۵ - سیگنال صدای قلب جوجه سالم در حوزه زمان-فرکانس و بر اساس لایه دوم ضرایب موجک گسسته



شکل ۶ - سیگنال صدای قلب جوجه بیمار در حوزه زمان-فرکانس و بر اساس لایه دوم ضرایب موجک گسسته

۲-۳ نتایج داده کاوی

در این مرحله، از تجزیه موجک هر سیگنال صدای قلب، ۶۹ ویژگی استخراج گردید. بیست و سه ویژگی (F1(AP2)-F23(AP2)) پارامترهای آماری ضرایب تقریب، بیست و سه ویژگی (F1(DE1)-F23(DE1)) پارامترهای آماری ضرایب جزئیات سطح یک و بیست و سه ویژگی (F1(DE2)-F23(DE2)) پارامترهای آماری ضرایب جزئیات سطح دوم می‌باشند. بر اساس روش IDE هر ویژگی امتیازدهی شد و داده‌هایی که بیشترین امتیاز را داشتند به عنوان بهترین ویژگی‌ها انتخاب شدند تا از آن‌ها به عنوان ورودی ماشین بردار پشتیبان استفاده گردد. شکل ۷ امتیاز ویژگی‌های استخراج شده از ضرایب موجک در سطح دوم را نشان می‌دهد. همان‌گونه که از شکل ۷ پیداست بین ویژگی ۶۵ و ۵۷ از لحاظ مقدار امتیاز تفاوت معناداری وجود دارد. به همین منظور حدآستانه امتیاز ۰/۷ در نظر گرفته شد و ۱۲ ویژگی نشان داده شده در شکل به عنوان ویژگی‌های برتر انتخاب شدند.



شکل ۷- ویژگی‌های حوزه زمان-فرکانس و امتیاز آن‌ها

۳-۳ دقت ماشین بردار پشتیبان

ویژگی‌های انتخاب شده در مرحله انتخاب ویژگی به عنوان ورودی SVM مورد استفاده قرار گرفتند. داده‌های پژوهش شامل ۲۱۴ نمونه سیگنال صدای قلب جوجه سالم و جوجه بیمار بود که ۸۰٪ نمونه‌ها (۱۷۲ نمونه سیگنال صدای قلب) به صورت کاملاً تصادفی برای آموزش SVM انتخاب گردیدند. ۲۰٪ باقی‌مانده (۴۲ نمونه سیگنال صدای قلب) به منظور آزمون ماشین بردار پشتیبان به کار گرفته شدند تا دقت سامانه طراحی شده به منظور تشخیص بیماری نیوکاسل مرغ بدست آید.

جدول ۱ معرف عملکرد SVM در روز ۱۶ (دو روز بعد از اعمال بیماری) بر اساس بهترین ویژگی‌های انتخاب شده توسط IDE از ضرایب موجک در سطح دوم می‌باشد. جداول ۲ و ۳ به ترتیب ماتریس‌های اغتشاش داده‌های آموزش و آزمون را برای حوزه زمان-فرکانس نشان می‌دهند.

با توجه به جدول ۱ بیشینه دقت ماشین بردار پشتیبان به ازای $\sigma = 0/8$ آمده است. مقدار σ معرف پهنای ابر صفحه می‌باشد. هر چقدر مقدار σ بیشتر باشد، ابر صفحه رفتار عمومی‌تری از خود نشان می‌دهد و با کم شدن مقدار σ ، ابر صفحه به سمت محلی شدن سوق پیدا می‌کند (Burges, 1998).

جدول ۱- ماتریس اغتشاش ماشین بردار پشتیبان برای بهترین ویژگی‌های انتخاب شده توسط IDE در سطح دوم ضرایب موجک

تعداد ویژگی‌های ورودی	تابع اصلی	σ	دقت SVM	
			آموزش	آزمون
۱۲	RBF	۳	٪ ۹۱/۲۸	٪ ۹۲/۸۶
		۲	٪ ۹۲/۴۴	٪ ۹۲/۸۶
		۱	٪ ۹۹/۴۲	٪ ۹۲/۸۶
		۰/۸	٪ ۱۰۰	٪ ۹۲/۸۶
		۰/۶	٪ ۱۰۰	٪ ۹۰/۴۸
		۰/۴	٪ ۱۰۰	٪ ۸۵/۷۱



از جدول ۱ مشخص است که SVM توانست با دقت ۹۲/۸۶٪ در روز دوم بیماری، جوجه بیمار را از جوجه سالم بر اساس سیگنال صدای قلبشان جدا کند.

جدول ۲، ماتریس اغتشاش داده‌های آموزش را نشان می‌دهد. با توجه به این جدول، SVM همه جوجه‌های سالم و بیمار را درست تشخیص داده است و با دقت ۱۰۰٪ به شناسایی آن‌ها پرداخته است. جدول ۳، ماتریس اغتشاش داده‌های آزمون را نشان می‌دهد. این جدول به وضوح بیان می‌دارد که SVM از ۲۱ عدد جوجه بیمار، با دقت ۱۰۰٪ همه را به درستی تشخیص داده است و از تعداد ۲۱ جوجه سالم فقط ۳ عدد را به اشتباه بیمار تشخیص داده است و در کل با دقت میانگین ۹۲/۸۶٪ به شناسایی و تفکیک جوجه‌های بیمار از سالم بر اساس سیگنال صدای قلبشان پرداخته است.

جدول ۲- ماتریس اغتشاش داده‌های آموزش حوزه زمان-فرکانس جدول ۳- ماتریس اغتشاش داده‌های آزمون حوزه زمان-فرکانس

		حوزه زمان- فرکانس		
		سالم	بیمار	دقت (%)
$\sigma = 0/8$	سالم	۱۸	۳	٪ ۸۵/۷۱
	بیمار	۰	۲۱	٪ ۱۰۰
		دقت میانگین = ٪ ۹۲/۸۶		

		حوزه زمان- فرکانس		
		سالم	بیمار	دقت (%)
$\sigma = 0/8$	سالم	۸۶	۰	٪ ۱۰۰
	بیمار	۰	۸۶	٪ ۱۰۰
		دقت میانگین = ٪ ۱۰۰		

۴- نتیجه گیری

هدف از این پژوهش طراحی سامانه‌ای هوشمند به منظور تشخیص بیماری نیوکاسل مرغ بر اساس سیگنال صدای قلب مرغ بود. صدای قلب توسط گوشی پزشکی ضبط شد و سیگنال‌های صدا توسط نرم افزار متلب مورد تحلیل قرار گرفتند. این سیگنال‌ها در حوزه زمان-فرکانس با استفاده از تبدیل موجک دابچی نوع اول و در سطح دوم پردازش شدند. در مرحله داده‌کاوی، ۶۹ ویژگی از ضرایب تقریب و جزییات موجک لایه دوم برای ۲۱۴ نمونه استخراج شدند و با استفاده از روش IDE، ۱۲ عدد از بهترین ویژگی‌ها انتخاب گردیدند. از این ویژگی‌ها به عنوان ورودی ماشین بردار پشتیبان بهره گرفته شد. دقت SVM در روز ۱۶ (دو روز بعد از اعمال بیماری) در مرحله آموزش و آزمون به ترتیب ۱۰۰٪ و ۹۲/۸۶٪ به دست آمد. نتایج این پژوهش کارا بودن روش‌های هوشمند را در تشخیص بیماری پرند از روی صدای قلب بیان می‌دارند.

منابع و مأخذ

- ۱- علی‌زاده، س.م. رفیعی، ع. نخعی، توکلی گلپایگانی، ع. ۱۳۹۲. تشخیص امراض دریچه‌ای قلب با استفاده از سیگنال PCG و شبکه عصبی، کنفرانس ملی مهندسی کامپیوتر و فن‌آوری اطلاعات.
- ۲- نخعی، ن. رفیعی، ع. علیزاده، س.م. ۱۳۹۲. طبقه بندی صدای قلب با استفاده از تبدیل ویولت و طرح خودسازماندهی افزایشی، همایش ملی مهندسی برق و توسعه پایدار با محوریت دستاوردهای نوین در مهندسی برق.



3. Acevedo, M. A. Corrada-Bravo, C. J. Corrada-Bravo, H. Villanueva-Rivera, L. J. & Aide, T. M. 2009. Automated classification of bird and amphibian calls using machine learning: A comparison of methods. *Ecological Informatics*, 4(4), 206-214 .
4. Bagheri, B. Ahmadi, H. & Labbafi, R. 2010. *Application of data mining and feature extraction on intelligent fault diagnosis by artificial neural network and k-nearest neighbor*. Paper presented at the XIX International Conference on Electrical Machines - ICEM , Rome, 2010.
5. Burges, C. J. 1998. A tutorial on support vector machines for pattern recognition. *Data mining and knowledge discovery*, 2(2), 121-167 .
6. Cortes, C. & Vapnik, V. 1995. Support-vector networks. *Machine learning*, 20(3), 273-297 .
7. Duhamel, P. & Vetterli, M. 1990. Fast Fourier transforms: a tutorial review and a state of the art. *Signal processing*, 19(4), 259-299 .
8. Iyer, S. Sinha, S. K. Tittmann, B. R. & Pedrick, M. K. 2012. Ultrasonic signal processing methods for detection of defects in concrete pipes. *Automation in Construction*, 22, 135-148 .
9. Lee, J. J. Lee, S. M. Kim, I. Y. Min, H. K. & Hong, S. H. 1999. *Comparison between short time Fourier and wavelet transform for feature extraction of heart sound*. Paper presented at the TENCON 99. Proceedings of the IEEE Region 10 Conference.
10. Lei, Y. He, Z. & Zi, Y. 2008. A new approach to intelligent fault diagnosis of rotating machinery. *Expert Systems with Applications*, 35(4), 1593-1600 .
11. Marchant, B. 2003. Time–frequency analysis for biosystems engineering. *Biosystems engineering*, 85(3), 261-281 .
12. Misiti, M. Misiti, Y. Oppenheim, G. & Poggi, J.-M. 1996. Wavelet toolbox. *The MathWorks Inc., Natick, MA* .
13. Morgan, H. 1946. Newcastle Disease Of Poultry. *Iowa State University Veterinarian*, 9(1), 4 .
14. Nidzworski, D. Wasilewska, E. Smietanka, K. Szewczyk, B. & Minta, Z. 2013. Detection and differentiation of Newcastle disease virus and influenza virus by using duplex real-time PCR. *Acta Biochimica Polonica*, 60(3), 475-480 .
15. Saravanan, N. & Ramachandran, K. 2010. Incipient gear box fault diagnosis using discrete wavelet transform (DWT) for feature extraction and classification using artificial neural network (ANN). *Expert Systems with Applications*, 37(6), 4168-4181 .
16. Scholkopf, B. Sung, K.-K. Burges, C. J. Girosi, F. Niyogi, P. Poggio, T. & Vapnik, V. 1997. Comparing support vector machines with Gaussian kernels to radial basis function classifiers. *Signal Processing, IEEE Transactions on*, 45(11), 2758-2765 .
17. Turkoglu, I. Arslan, A. & Ilkay, E. 2003. An intelligent system for diagnosis of the heart valve diseases with wavelet packet neural networks. *Computers in Biology and Medicine*, 33(4), 319-331 .
18. Wang, X. & Makis, V. 2009. Autoregressive model-based gear shaft fault diagnosis using the Kolmogorov–Smirnov test. *Journal of Sound and Vibration*, 327(3), 413-423 .



نهمین کنگره ملی مهندسی ماشین‌های کشاورزی

(مکانیک بیوسیستم) و مکانیزاسیون

پردیس کشاورزی و منابع طبیعی دانشگاه تهران

۲ و ۳ اردیبهشت ۱۳۹۴ - کرج



Newcastle disease diagnosis of chickens based on wavelet analysis of heart's sound

Abstract

In this research, Newcastle disease was diagnosed based on their heart's sound. In the first, the birds were divided into two categories. One category was infected by Newcastle disease virus and their heart's sounds were recorded by stethoscope. Their heart's sounds were transmitted to the time- frequency domain using Discrete Wavelet Transform (DWT). In the next step, 69 features were extracted from approximates and details coefficients of Dubechi 1 and two levels of wavelet transform. In order to reach to the best feature, 12 features were selected by using IDE method. Support vector machine (SVM) and radial basic function (RBF) were used as a classifier and separator respectively. At the first step of classification, in order to reach to the best structure of classifier, SVM has been trained by 172 heart's signal and after that has been analyzed by 42 heart's signal. SVM accuracy in diagnosis of Newcastle disease based on heart's sound on 2th day after the disease, has been obtained 92.86 %.

Keywords: Newcastle disease, heart's Sound, Support vector machine, Data mining, Wavelet discrete transform