پژوهنده (مجله پژوهشی دانشگاه علوم پزشکی شهید بهشتی)
سال نوزدهم، شماره ۱، پی در پی ۹۷، صفحات ۴۸ تا ۵۴
فروردین و اردیبهشت ۱۳۹۳

تاریخ دریافت مقاله: ۱۳۹۲/۲/۱
تاریخ پذیرش مقاله: ۱۳۹۲/۱۲/۲۴

طراحی دادگان اعمال جراحی ترمیمی دست جهت سامانه تبدیل گفتار به متن

*دکتر مرجان قاضی‌سعیدی[۱]، دکتر رضا صفدری[۲]، دکتر عبدالجلیل کلانتر هرمزی[۳]، دکتر لیلا شاهمرادی[۴]، فاطمه صادقی[*۵]*

۱. دکتری مدیریت اطلاعات سلامت، مربی گروه مدیریت اطلاعات سلامت، دانشکده پیراپزشکی، دانشگاه علوم پزشکی تهران
۲. دکتری مدیریت اطلاعات سلامت، دانشیار گروه مدیریت اطلاعات سلامت، دانشکده پیراپزشکی، دانشگاه علوم پزشکی تهران
۳. استاد گروه جراحی پلاستیک و ترمیمی، دانشکده پزشکی، دانشگاه علوم پزشکی شهید بهشتی
۴. دکتری مدیریت اطلاعات سلامت، مربی گروه مدیریت اطلاعات سلامت، دانشکده پیراپزشکی، دانشگاه علوم پزشکی تهران
۵. کارشناس ارشد آموزش مدارک پزشکی، دانشکده پیراپزشکی، دانشگاه علوم پزشکی تهران

چکیده

سابقه و هدف: نرم‌افزار تبدیل گفتار به متن، ابزار تسهیل‌کننده‌ای جهت ورود داده به پرونده‌ی الکترونیک سلامت می‌باشد. بکارگیری این نرم‌افزار در حوزه‌ی بهداشت و درمان، مستلزم تعریف دادگان تخصصی جهت آن است. هدف پژوهش حاضر، ایجاد دادگان متنی تخصصی جهت نرم‌افزار تبدیل گفتار به متن در جراحی ترمیمی دست و تست عملی این فناوری در بیمارستان می‌باشد.

مواد و روش‌ها: مرحله‌ی اول پژوهش به روش اکتشافی برای طراحی دادگان و مرحله‌ی دوم جهت تعیین کارایی سیستم به روش تجربی انجام شد. مرحله اول، شامل تایپ ۱۸۶۳ برگ شرح عمل مربوط به دوره‌ی ۳ ماهه‌ی اول سال ۱۳۹۱ در بیمارستان ۱۵ خرداد می‌باشد. پس از انجام تحلیل‌های موردنیاز، دادگان متنی جراحی ترمیمی دست ایجاد و بر روی سیستم نویسا، بارگذاری شد. در مرحله‌ی دوم جهت تعیین میزان دقت نرم‌افزار، سیستم در محیط اتاق عمل تست و ارزیابی شد.

یافته‌ها: طراحی دادگان شامل سه مرحله‌ی ایجاد متون الکترونیکی، استخراج مدل زبانی و واژگان از این متون می‌باشد. دادگان ایجاد شده، شامل واژگان بزرگ به مفهوم بیش از ۱۰۰۰ واژه، جهت بکارگیری در واحد جراحی دست در بیمارستان‌های مربوطه است. پس از تست نرم‌افزار در اتاق عمل دقت سیستم برای کاربر آموزش دیده ۸۳/۷٪ و در کاربر بدون آموزش، ۷۷/۸٪ به دست آمد.

نتیجه‌گیری: این پروژه منتج به ایجاد نرم‌افزاری در حوزه‌ی پزشکی منطبق با شرایط کشور شد که امکان ایجاد مدارکی کامل و الکترونیکی را فراهم خواهد کرد. دقت سیستم با صداهای زمینه، آموزش سیستم با صدای کاربر و نوع میکروفن مرتبط می‌باشد، به طوری‌که بهبود هر یک از موارد فوق در زمان تست، دقت سیستم را افزایش خواهد داد.

واژگان کلیدی: دادگان، جراحی ترمیمی دست، نرم‌افزار تبدیل گفتار به متن

لطفاً به این مقاله به صورت زیر استناد نمایید:

Ghazisaeedi M, Safdari R, Kalantar-Hormozi A, Shahmoradi L, Sadeghi F. The design of a reconstructive hand surgery text database based on a speech recognition system. Pejouhandeh 2014;19(1):48–54.

مقدمه

مستندسازی ناقص پرونده‌های پزشکی توسط پزشکان، علاوه بر از دست رفتن اطلاعات بیماران، اثرات سویی در فرایند درمانی داشته (۱) و هزینه‌های زیادی را به سیستم تحمیل می‌کند (۲). کلینیک‌ها، بیمارستان‌ها و سایر مؤسسات مراقبت بهداشتی، هر هفته از صدها و حتی هزاران بیمار مراقبت می‌کنند که این امر منجر به تولید حیرت‌آور پرونده‌ها می‌شود (۳) و در صورت استفاده از کاغذ با مسأله کمبود فضا جهت بایگانی و صرف زمان بسیار جهت بازیابی این پرونده‌ها مواجه می‌شویم. به نظر می‌رسد، عدم ثبت کامپیوتری اطلاعات یکی از دلایل نقص مدارک‌پزشکی در مراکز درمانی می‌باشد (۴). یکی از اوراق اصلی پرونده، برگ گزارش عمل جراحی است، که اقدامات جراحی انجام شده در طول اقامت بیمار در بیمارستان را منعکس می‌نماید (۵).

ثبت اقدامات درمانی و جراحی در پیگیری مراقبت و بهبود بیمار تأثیر به سزایی دارد اما کمبود وقت پزشکان و تعداد زیاد بیماران اورژانسی در این حوزه، موجب بی‌توجهی و نقص در ثبت پرونده‌ها می‌شود. بکارگیری فن‌آوری اطلاعات سلامت به پردازش اطلاعات، ذخیره، بازیابی، تسهیم و استفاده از

[*] **نویسنده مسؤول مکاتبات: فاطمه صادقی**؛ دانشکده پیراپزشکی، دانشگاه علوم پزشکی تهران، تلفن: ۰۹۱۲۵۹۰۰۶۱۲؛ **پست الکترونیک:** Fateme_s2860@yahoo.com

[Downloaded from pajoohande.sbmu.ac.ir on 2026-11-06]

اطلاعات مراقبت بهداشتی، داده‌های پزشکی و دانش برای ارتباطات و تصمیم‌گیری، کمک می‌کند (۶).

یکی از مهمترین اعضای بدن انسان از نظر ظاهر و عملکردی دست می‌باشد که حساسیت آن، انسان را از صدمات مصون نگه داشته و در غیاب حس بینایی، باعث شناخت اجسام می‌شود. از طرفی یک سوم صدمات وارده به انسان، دست را نیز گرفتار می‌کند (۷ و ۸). در بیشتر موارد، عدم توانایی فرد در انجام فعالیت‌های شغلی و روزمره می‌تواند به بیمار از نظر روانی آسیب برساند. لذا هر اقدام جهت برگرداندن سلامت دست بیمار، می‌تواند بر روی زندگی فرد از نظر فعالیت یا سلامت روانی ارزشمند باشد (۹).

با وجود تمایل اغلب کارکنان بالینی نسبت به استفاده از سیستم‌های اطلاعاتی، بیشتر این کارکنان برای ورود داده‌ها به کامپیوتر با چالش مواجه بوده و زمان صرف شده آنها از زمان مورد نیاز برای دست‌نویس کردن یا دیکته کردن، بیشتر می‌باشد (۱۰). به همین علت، تعبیه‌ی راه‌حلی که بدون نیاز به اطلاعات کامپیوتری یا زمان به مستندسازی کمک کند، موجب تکمیل پرونده و در نهایت ارتقای مراقبت می‌شود. اخیراً تعدادی فناوری جهت مستندسازی بالینی و سهولت تکمیل پرونده از قبیل سیستم بارکد، اسکن، قلم نوری، دستیار دیجیتال شخصی و سیستم تبدیل گفتار به متن (۱۱) ایجاد شده‌اند که از این میان سیستم تبدیل گفتار به متن در سال‌های اخیر به سرعت توسعه یافته است (۱۲).

اولین گزارش‌ها از بکارگیری این فناوری در پزشکی مربوط به واحد رادیولوژی است (۱۳ و ۱۴). بر طبق بررسی‌های انجام شده در کشورهایی مثل آمریکا و استرالیا، شرکت‌های بزرگ تجاری نانس و آگفا، نرم‌افزار را توسعه و عرضه نمودند به طوری که اغلب نرم‌افزارهای موجود، از محصولات این شرکت‌ها می‌باشد (۱۸-۱۵).

از آن جمله می‌توان به نسخه‌ی پزشکی دراگون (Dragon) اشاره نمود که به عنوان راه حل بخش بالینی جهت فراهم نمودن ورود سریع‌تر و راحت‌تر اطلاعات به درون پرونده‌ی الکترونیک سلامت Electronic Health Record (EHR) ایجاد شده است. با این نرم‌افزار می‌توان به پیاده‌سازی و پیشرفت EHR اطمینان داشت (۱۹).

هدف تحقیقات اخیر، ارتقا و پیشرفت سامانه‌ای کاملاً دقیق یا با حداقل درصد خطا در حوزه‌های مختلف علمی از جمله بهداشت و درمان است (۲۰). پایه‌ی کار سیستم تبدیل گفتار به متن، تعریف دادگان مربوط به آن گفتار در نرم‌افزار درونی سیستم می‌باشد (۲۱). پژوهش حاضر به ایجاد دادگان متنی تخصصی جهت نرم‌افزار تبدیل گفتار به متن نویسا، در حوزه‌ی جراحی ترمیمی دست پرداخته و پس از اتمام مرحله‌ی طراحی، تست عملی فناوری در بیمارستان ۱۵ خرداد انجام شده است تا قابلیت پیاده‌سازی سیستم را در حوزه‌ی بهداشت و درمان بسنجد.

مواد و روش‌ها

مرحله‌ی اول پژوهش، به روش اکتشافی برای طراحی دادگان و مرحله‌ی دوم، جهت تعیین کارایی آن به روش تجربی انجام شد. در این مطالعه، کلیه‌ی شرح عمل‌های بیماران در حیطه‌ی اعمال جراحی ترمیمی دست در ۳ ماهه‌ی اول سال ۱۳۹۱ (اول فروردین ماه الی ۳۱ خرداد) در بیمارستان ۱۵ خرداد وابسته به دانشگاه علوم پزشکی شهید بهشتی جمع‌آوری شد. در مجموع، ۱۸۶۳ برگ شرح عمل، برای مرحله‌ی طراحی دادگان به دست آمد که تمام شرح‌عمل‌ها تایپ و تبدیل به متون الکترونیکی گردید. از این بین، تعداد ۱۰۸ شرح عمل به طور تصادفی انتخاب و در طی سه ماه اول جمع‌آوری داده، از پزشکان (۵ رزیدنت سال اول جراحی پلاستیک و ترمیمی و یک نفر فلوشیپ جراحی دست) درخواست شد که پس از نوشتن شرح عمل، فعالیت انجام شده را به طور شفاهی برای دستگاه Player Mp3 دیکته نمایند. سپس در فایل جداگانه‌ای، ۱۰۸ شرح‌عمل ضبط شده، به متون الکترونیکی تبدیل و با ۱۰۸ شرح عمل کاغذی مشابه، مورد مقایسه قرار گرفتند. این بخش برای دستیابی به تفاوت میان کلمات دیکته شده با کلمه نگارش شده همچنین چگونگی تلفظ کلمات توسط پزشکان انجام شده است.

پس از ایجاد متون الکترونیکی، نیاز به تولید مدل زبانی بود که این بخش شامل تهیه‌ی لیستی از کلیه‌ی کلمات متن، همراه با تعداد تکرار هر کلمه در آن بود. در این مورد کل فایل الکترونیکی به صورت کلمه‌ای تفکیک، سپس با استفاده از نرم‌افزار اکسل (Excell) در بخش اول بسامد تکرار هر واژه و در بخش دوم، بسامد تکرار سه واژه با هم جهت ایجاد مدل زبانی ترایگرام (Trigram) محاسبه گردید.

تعریف واژگان برای نرم‌افزار از طریق نگارش تلفظ هر واژه با استفاده از استاندارد الفبای آوانگاری بین‌المللی International Phonetic Alphabet (IPA) صورت می‌گیرد. در این الفبا، برای نشان دادن هر آوا، از یک نویسه‌ی خاص استفاده می‌شود و رابطه‌ای یک به یک میان حرف و آوا در آن وجود دارد به نحوی که به کاربران اجازه می‌دهد تا آواها، گویش‌ها و لحن‌های زبان‌های مختلف را به درستی و به

[Downloaded from pajoohande.sbmu.ac.ir on 2026-11-06]

طور جامع بنویسند. این استاندارد در کشور ایران شامل ۲۹ واج است که برای بخش تلفظ یا آوانگاری سیستم مورد استفاده قرار گرفت (جدول ۱).

جدول ۱. ۲۹ واج مربوط به الفبای تلفظ‌نویسی یا قراردادی از الفبای IPA در ایران

U	I	AA	A	E
خ	آی	آ	أ	ای
O	Y	SH	S	R
أ	ی	ش	س	ر
X	ZH	Z	L	K
خ	ژ	ز	ل	ک
M	N	B	V	CH
م	ن	ب	و	چ
P	F	D	AH	T
پ	ف	د	ع	ت
JE	H	G	Q	
چ	ح	گ	ق	

برای رعایت استانداردهای مدارک پزشکی و تایپ نرم‌افزار در محیط شرح عمل استاندارد، این برگ در نرم‌افزار مایکروسافت ورد (Microsoft Word) طراحی و برای استفاده در نرم‌افزار تعریف گردید.

در مرحله‌ی تست سیستم، پس از نصب نرم‌افزار و هماهنگی با ۶ نفر رزیدنت سال اول جراحی پلاستیک جهت دیکته‌ی شرح عمل، پزشکان به صورت تصادفی به ۲ گروه تقسیم شدند. گروه اول به مدت ۲۰ دقیقه تحت آموزش روش خواندن و تطبیق صدا با نرم‌افزار قرار گرفتند، در حالی که گروه دوم بدون گذراندن دوره‌ی آموزش با نرم‌افزار، کار دیکته‌ی شرح عمل را انجام دادند. به ازای هر پزشک، ۱۰ برگ شرح عمل در طی دو هفته زمان تست، انتخاب و برای سیستم دیکته و تبدیل به متن گردید. سپس این ۶۰ فایل دیکته شده، مورد تجزیه و تحلیل قرار گرفت تا میزان دقت سیستم در مرحله‌ی تست مشخص گردد. برای تحلیل و آنالیز میزان دقت و خطای سیستم در بازشناسی گفتار، از فرمول‌های ۱ و ۲ مطابق زیر استفاده گردید:

$$WER\% = 1 - Accuracy\% = \frac{N^{Del.} + N^{Ins.} + N^{Sub.}}{N^{All}} \times 100$$

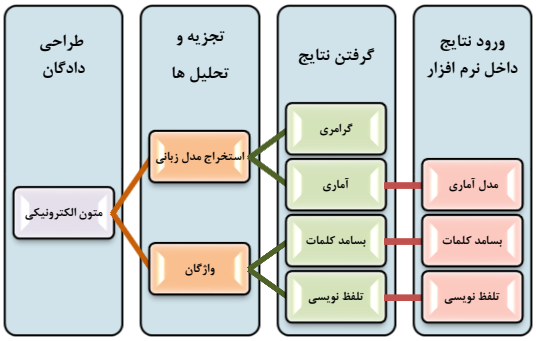
فرمول ۱. نرخ خطای کلمات (Word Error Rate). $N^{Del.}$: تعداد کلمات حذف؛ $N^{Ins.}$: تعداد کلمات درج؛ $N^{Sub.}$: تعداد کلمات نبود دادگان؛ N^{All}: تعداد کل کلمات دیکته شده.

$$Accuracy\% = \frac{N^{All} - (N^{Del.} + N^{Ins.} + N^{Sub.})}{N^{All}} \times 100$$

فرمول ۲. درصد دقت. $N^{Del.}$: تعداد کلمات حذف؛ $N^{Ins.}$: تعداد کلمات درج؛ $N^{Sub.}$: تعداد کلمات نبود دادگان؛ N^{All}: تعداد کل کلمات دیکته شده.

یافته‌ها

طراحی دادگان شامل چند مرحله است: اولین مرحله، ایجاد متون الکترونیکی یا پیکره‌ی متنی حوزه‌ی مورد نظر می‌باشد که از این متون، موارد مربوطه استخراج و به بخش آماده دادگان گفتاری اضافه می‌شود تا نرم‌افزار، قابل بکارگیری در تخصص مورد نظر گردد. این مراحل در شکل ۱، نشان داده شده است.



شکل ۱. مراحل و اجزای طراحی دادگان متنی در سیستم تبدیل گفتار به متن.

جهت یکدست‌سازی کلمات پیکره‌ی متنی، نگارش‌های متفاوت کلمات شرح عمل موجود در دو فایل متنی و صوتی (ضبط شده)، با یکدیگر مقایسه گردید. برای مثال یک کلمه با ۳۳ بار تکرار در فایل صوتی و ۳۳ بار تکرار در فایل متنی انتخاب و مقایسه شد که بیانگر تفاوت فایل‌های دیکته شده توسط پزشکان با فایل‌های حاصل از تایپ همان برگ شرح‌عمل می‌باشد. یافته‌های حاصل از این مقایسه در جدول شماره ۱، بر اساس مقایسه‌ی انواع کلمات موجود در دو متن، آورده شده است. این مقایسه، نیاز به یکدست‌سازی متون را به درستی نمایش می‌دهد.

بر طبق یافته‌های حاصل از جدول ۱، رعایت نشدن یک استاندارد خاص در نگارش متون شرح عمل، در محیط رایانه‌ای مشخص شده و باعث ایجاد کلمات یکسان با کاراکترهای متفاوت می‌گردد و هنگام شمردن این کلمات، به عنوان دو کلمه‌ی مختلف در نظر گرفته می‌شوند. چنین مواردی در اصل یک تلفظ یکسان برای چند شکل نوشتاری ایجاد کرده و موجب بروز خطای آکوستیکی در نرم‌افزار می‌شوند. بنابراین، اطلاعات موجود باید به قالبی تبدیل شوند که استخراج اطلاعات آماری از آنها به راحتی امکان‌پذیر باشد. به همین دلیل، حالات املایی مختلف کلمات جستجو و به یک حالت واحد تبدیل گردید.

برای بخش نگارش تلفظ، علاوه بر ثبت تلفظ‌های متفاوت،

[Downloaded from pajoohande.sbmu.ac.ir on 2026-11-06]

صورت‌های نوشتاری نیز تحت عنوان تلفظ لحاظ گردید و به تلفظ کلمه اضافه شد. تعداد کل کلمات متن در بخش تلفظ‌نویسی (بعد از یکدست‌سازی) ۱۳۲۴ کلمه و تعداد تلفظ از یک مورد تا ۱۸ مورد در انواع کلمات متفاوت بود. در کل متن فقط دو کلمه با تلفظ یکسان و معنی متفاوت وجود داشت. این کلمات شامل تلفظ PIN (به دو معنی ابزاری برای فیکس شکستگی و خردشدگی استخوانی و مورد دوم حالت اختصاری مربوط به عصب Posterior Interosseous Nerve) و تلفظ "IN" در دو کلمه‌ی "این" و "in" است.

جدول ۲ : مقایسه انواع کلمات موجود در دو متن و تطبیق نمونه‌های متفاوت.

کلمه درون فایل‌ها	تعداد کلمه در فایل صوتی	تعداد کلمه در فایل متنی
Flap	موجود نیست	۵
فلاپ	۱۵	۱۱
فلپ	۱۸	۱۷
جمع	۳۳	۳۳

در مرحله‌ی تست سیستم، بر حسب اینکه پزشکان آموزش کاربردی دیده باشند یا خیر، ۶۰ پرونده بین ۶ پزشک در اتاق عمل دیکته شد. سپس میزان شاخص‌های خطا تعیین گردید (جدول ۳). بررسی‌های انجام شده نشان داد که کاربران آموزش ندیده با تعداد ۸۴۸ کلمه‌ی دیکته شده، حدود ۱۸۱ کلمه خطا داشته‌اند که فراوانی ۲۱/۳٪ وجود داشت و در گروه آموزش دیده، با ۵۵۷ کلمه، که ۹۰ کلمه خطا داشتند، فراوانی ۱۶/۱٪ به دست آمد. ارتباط آموزش بین این دو گروه با P<۰/۰۵ نشان‌دهنده‌ی معنی‌دار بودن ارتباط و تأثیر آموزش در بهبود عملکرد نرم‌افزار می‌باشد.

در مرحله‌ی تست سیستم تبدیل گفتار به متن، یکی از سه نوع خطای حذف، درج و جایگزینی ممکن است رخ دهد. خطای حذف، شامل وجود یک واحد آوایی (کلمه یا واج) در سیگنال گفتار می‌باشد، ولی بازشناسی نمی‌شود. خطای درج، واحد آوایی بازشناسی شده در سیگنال گفتار وجود ندارد که معمولاً در هنگام تشخیص نویز به جای یک واحد آوایی پیش می‌آید. وقتی که یک واحد آوایی به اشتباه به جای یک واحد آوایی دیگری بازشناسی می‌شود خطای جایگزینی رخ داده است (۲۲).

همان‌طور که در جدول ۲ نشان داده شده است، میزان خطای حاصل از کلماتی که در دادگان تعبیه نشده، ۳٪ می‌باشد. البته این کلمات، به راحتی قابلیت افزایش به دادگان را داشته و موجب غنی‌تر شدن دادگان در مراحل بعدی اجرای پروژه و بهبود دقت تشخیص سیستم می‌شود. مقایسه‌ی کاربران با میزان آموزش متفاوت برای استفاده درست سیستم و بکارگیری جهت دیکته‌ی شرح عمل نیز با P-value<۰/۰۵ بیانگر معنی‌دار بودن ارتباط آموزش کاربران با بهبود دقت سیستم می‌باشد.

مهمترین خطای ایجاد شده در این آزمایش، مربوط به نویز محیط و حدوداً ۱۳/۷٪ می‌باشد که به عنوان علت اصلی میزان خطا در سیستم و درج اشتباه کلمه است. در نتیجه، مهمترین چالش پیش روی پژوهشگر، کاهش این خطا با حذف نویز محیط می‌باشد.

جدول ۳. میزان و نوع خطاهای موجود در نرم‌افزار بعد از تست سیستم در اتاق عمل.

خطاها	کاربر آموزش دیده تعداد (درصد)	کاربر بدون آموزش تعداد (درصد)
کلمات خارج از دادگان	۱۰(۱/۸)	۳۲(۳/۸)
خطای جایگزینی	۳(۰/۵)	۹(۱/۱)
خطای دیکته کننده	۳(۰/۵)	۶(۰/۷)
خطای درج (نویز)	۶۹(۱۲/۴)	۱۲۴(۱۴/۶)
خطای حذف	۵(۱)	۱۰(۱/۱)
کل خطا	۹۰(۱۶/۲)	۱۸۱(۲۱/۳)

بحث:

پژوهش نشان داد که طراحی دادگان، مقدور بوده و دارای کارایی لازم برای اعمال جراحی ترمیمی دست می‌باشد. نرم‌افزار تشخیص صدا امواج را براساس سه معیار دادگان، پروفایل فرد دیکته‌کننده و مدل زبانی بازشناسی می‌نماید (۲۳).

دادگان سیستم‌های تبدیل گفتار به متن، به عنوان زیربنای اصلی سیستم، باید از جامعیت کافی برخوردار باشند. دادگان مناسب دارای حداکثر واژگان مربوط به تخصص مورد نظر، با قابلیت افزایش اصطلاحات جدید است (۲۳). بر اساس موارد فوق، در پژوهش سعی شد بیشترین تعداد شرح عمل برای طراحی دادگان به کار رود تا دادگان تولید شده، نمونه‌ی کاملی از کلمات تخصصی را دارا باشند.

بر اساس پژوهش Kirriemuire، در قلب نرم‌افزار، بخش ترجمه‌ی کلمات گفتاری به صدای پایه‌ای، شامل کلمات و سیلاب‌ها قرار دارد. برای بکارگیری نرم‌افزار در تخصص خاص نیاز به ایجاد واژگان تخصصی است تا در مرحله‌ی ترجمه، موجب افزایش دقت سیستم در آن حوزه گردد (۲۴). پژوهش ما در راستای پژوهش Kirriemuire می‌باشد.

[Downloaded from pajoohande.sbmu.ac.ir on 2026-11-06]

دادگان سیستم‌های بازشناسی گفتار بر اساس تعداد کلمات دادگان، به چهار دسته شامل واژگان کوچک ۱ تا ۱۰۰ کلمه، واژگان متوسط بین ۱۰۱ تا ۱۰۰۰ کلمه، واژگان بزرگ بین ۱۰۰۱ تا ۶۰۰۰۰ کلمه و واژگان خیلی بزرگ، بیشتر از ۶۰۰۰۰ کلمه تقسیم می‌شوند (۲۵). بر این اساس دادگان ایجاد شده در پژوهش شامل دادگان بزرگ به مفهوم بیش از ۱۰۰۰ واژه می‌باشد.

در پژوهش Souto بیان می‌شود که نوع متون برای ایجاد مدل زبانی، بستگی به حوزه‌ی کاربرد آن در بازشناسی گفتار دارد. جهت ایجاد مدل زبانی در تشخیص گزارش‌های پزشکی، نیاز به تعداد زیادی گزارش‌های پزشکی یا متون پزشکی مرتبط می‌باشد تا مدل زبانی استخراج گردد. این پژوهش در راستای پژوهش Souto می‌باشد (۲۶). از طرفی، سوتو کلمات خارج از واژگان را به عنوان یکی از مشکلات در کاهش دقت سیستم بیان می‌نماید که شامل کلمات بیان شده در گفتار است که در دادگان وجود ندارد. ساده‌ترین راه حل این مسأله، گسترش اندازه واژگان برای کاهش خطای مربوط به کلمات خارج از واژگان می‌باشد (۲۶). با وجود بکارگیری تعداد زیاد شرح عمل در طراحی دادگان، هنوز خطای ناشی از کلمات خارج از دادگان در مرحله تست وجود داشت که با قابلیت افزایش کلمات دادگان، این مورد در طول مرحله تست مرتفع گردید. این مورد نیز هم‌راستای نتایج حاصل از پژوهش می‌باشد.

از چالش‌های اساسی پیش روی محققان حوزه‌ی بازشناسی گفتار، دقت سیستم‌های تبدیل گفتار به متن می‌باشد (۲۷). دقت سیستم به وسیله‌ی اندازه‌گیری درصد تشخیص کلمه ارزیابی می‌شود (۲۸) که به چندین ویژگی شامل محیط، صداهای زمینه، آموزش سیستم با صدای کاربر و نوع میکروفن بستگی دارد (۲۹ و ۳۰).

طبق بررسی Wagner و همکاران، اگر صداهای زمینه زیاد گردد، دقت سیستم ممکن است تا ۸۰٪ کاهش یابد (۳۱). همچنین، در محیطی با چند فرد دیکته کننده و پرسنلی که با تلفن و یا با هم صحبت می‌کنند، همچنین پزشکانی که به بیماران مشاوره می‌دهند، تعداد و درصد دقت در هنگام دیکته به مقدار زیادی کاهش یافته و خطاهای جملات، بیشتر می‌شوند (۳۲). همچنین، طبق تحقیقات Monlar، وجود اتاق کوچکی بدون پرسنل و صدا که یک سیستم فردی در آن تعبیه شده باشد (۳۳)، درصد دقت بیشتری در دیکته کردن را فراهم خواهد کرد. در مرحله‌ی تست، نبود محل بدون صدا و مشخص برای دیکته، به عنوان علت اصلی ایجاد خطای درج به حساب آمد که با تعبیه‌ی محیطی بدون صدا می‌توان این خطا را کاهش داد یا حتی حذف نمود. یافته‌های حاصل از پژوهش، تأییدکننده‌ی پژوهش‌های Monlar و Wanger بود.

آموزش کافی کاربران به عنوان موردی مهم برای موفقیت پیاده‌سازی‌های نرم‌افزار محسوب می‌شود (۳۲). همچنین، بر اساس پژوهش Zafar، آموزش کاربر برای صحبت مداوم و واضح‌تر، موجب کاهش میزان خطاها می‌گردد. Zafar بیان می‌دارد که گذراندن ۴۵ دقیقه آموزش با سیستم برای تشخیص صدا و آموزش فرد، جهت کار با سیستم، ضروری است (۳۴). این مورد، با توجه به نتایج حاصل از مرحله‌ی تست سیستم و کاربران با آموزش متفاوت در پژوهش فعلی نیز تأیید شد. میزان آموزش سیستم، به نرم‌افزار مورد استفاده، قدرت پردازشی کامپیوتر و میزان دقت مورد انتظار از سیستم، بستگی دارد. بکارگیری کامپیوترهای دارای قدرت پردازشی بالاتر، موجب کاهش میزان و زمان آموزش می‌شود (۲۴). این آموزش در سیستم فعلی، ۲۰ دقیقه است. همچنین در پژوهش Grasso نیز یافته‌ها حاکی از این است که آموزش و دقت سیستم با هم مرتبط هستند به طوریکه بعد از آموزش ۳۰ دقیقه برای هر فرد، عملکرد سیستم برای آن حوزه‌ی خاص، بهبود می‌یابد (۳۵). نتایج حاصل از پژوهش اخیر نیز هم‌راستای پژوهش Grasso می‌باشد. در پژوهشی توسط Jungk و همکاران با استفاده از سیستم IBM برای مستندسازی موارد بیهوشی، فاز آموزش نرم‌افزار، ۹۰ دقیقه بوده است که ۳۰ دقیقه‌ی آن جهت آشنایی روش صحبت با نرم‌افزار اعمال شده است. این میزان از آموزش با بهبود دقت سیستم همراه بود (۳۶). نتایج این پژوهش هم‌راستای پژوهش فوق می‌باشد.

این پروژه به عنوان کاری بنیادی در حوزه‌ی پزشکی قابلیت بکارگیری نرم‌افزار جهت بهبود و تکمیل پرونده‌ی الکترونیک را تأیید نمود. البته نیاز به چندین مرحله تست و بهینه‌سازی، جهت بهبود دقت سیستم و ارزیابی اجرای عملی آن، لازم می‌باشد. دادگان طراحی شده، شامل واژگان بزرگ به مفهوم بیش از ۱۰۰۰ واژه برای تخصص جراحی دست و سیستم تشخیص گفتار پیوسته بوده که قابلیت بکارگیری در بیمارستان‌های جراحی پلاستیک و ترمیمی در بخش جراحی دست را دارد. این نرم‌افزار، اولین تلاش برای معرفی سیستم تبدیل گفتار به متن در حوزه‌ی پزشکی است که با مستندسازی بهینه و جلوگیری از بروز خطا در ورود اطلاعات بیمار، موجب تکمیل اطلاعات پرونده می‌شود. نرم‌افزار حاضر از نوع بلادرنگ، مستقل از گوینده و مبتنی بر کلمه است. در

[Downloaded from pajoohande.sbmu.ac.ir on 2026-11-06]

این پروژه، از سیستم پایه "نویسا" و دادگان گفتاری این نرم‌افزار برای بکارگیری دادگان متنی تولیدی، استفاده شد.

تشکر و قدردانی

از مسؤولان و پزشکان بیمارستان فوق تخصصی ۱۵ خرداد به دلیل همکاری و مساعدت در ایجاد و تست پروژه، کمال تشکر را داریم. همچنین، از مدیر و مهندسین شرکت عصر گویش پرداز به دلیل قبول پیاده سازی دادگان بر روی نرم‌افزار نویسا و زحمات بی‌شائبه‌ی ایشان، صمیمانه قدردانی می‌شود.

REFERENCES

1. Mashoufi M, Rostami K, Mardi A. Documentation of medical records by physicians in the hospitals under Ardabil University of Medical Sciences 2001. J Ardebil Univ Med Sci 2006;6(1):73–7. (Full Text in Persian)

2. Afifyan M. Design of electronic system for classification of oral disease according to the WHO classification systems in 2010. Tehran: Tehran University of Medical Sciences; 2010. (Full Text in Persian)

3. Sedghi Jahromi M, Langarizade M. Introduction to health information technology. Tehran: Jafari; 2011. p. 204–56. (Full Text in Persian)

4. Furuzande A. Evaluation of the complete medical records in public hospitals of Shiraz University of Medical Sciences in 1998. Healthcare Inf Manage 2004;1(1):15–4. (Full Text in Persian)

5. Safdari R, Masouri N, Sieyedfarajalah S. Electronic health record system (standards, health data). Tehran: Jafari; 2011. p. 254–78. (Full Text in Persian)

6. Torabi M, Safdari R, Shahmoradi L. Health information technology management. Tehran: Jafari; 2010. p. 238–25. (Full Text in Persian)

7. Ahmadi A. Principles of hand surgery examination, diagnosis and treatment of common problems. Tehran: Moalef; 1992. p. 189–37. (Full Text in Persian)

8. Yavari M, Abdolrazagh HA, Riyahi A. Comparison of the results of single tendon transfer with ternary tendon transfer surgery in patients with radial nerve palsy. Pejouhandeh 2012;17(3):148–51. (Full Text in Persian)

9. Radmehr A. Survey results of gradual long way off bones of the fingers and palms in the hospital (Dissertation). Tehran: Shahid Beheshti University of Medical Sciences; 2008. (Full Text in Persian)

10. Sadoughi F, Ghazisaeid M, kimiafar KH, Ramezanghorbani N. Health information management technology. Tehran: Jafari; 2011. p. 259–52. (Full Text in Persian)

11. Yen PY, Gorman PN, editors. Usability testing of a digital pen and paper system in nursing documentation. AMIA Annual Symposium Proceedings. American Medical Informatics Association; 2005.

12. Monnich G, Wetter T. Requirements for speech recognition to support medical documentation. Methods Inf Med 2000;39(1):63–9.

13. Derman YD, Arenovich T, Strauss J. Speech recognition software and electronic psychiatric progress notes: physicians' ratings and preferences. BMC Med Inf Decis Making 2010;10(1):44–9.

14. Sas J, Poreba T. Optimal acoustic model complexity selection in polish medical speech recognition. Int J Med Inform 2011;17:115–22.

15. Rosenthal D, Chew F, Dupuy D, Kattapuram S, Palmer W, Yap R, *et al*. Computer-based speech recognition as a replacement for medical transcription. Am J Roentgenol 1998;17(1):23–5.

16. McGurk S, Brauer K, Macfarlane T, Duncan K. The effect of voice recognition software on comparative error rates in radiology reports. Br J Radiol 2008;81(970):767–70.

17. Pezzullo JA, Tung GA, Rogg JM, Davis LM, Brody JM, Mayo-Smith WW. Voice recognition dictation: radiologist as transcriptionist. J Digit Imaging 2008;21(4):384–9.

18. Issenman RM, Jaffer IH. Use of voice recognition software in an outpatient pediatric specialty practice. Pediatrics 2004;114(3e):290–3.

19. Nounce. Dragon Medical 360 Solutions. United States & Canada: Nuance Communications; 2011 [cited 2012 March 10]. Available from: http://www.nuance.com/for-healthcare/dragon-medical-360/index.htm.

20. Sufifar M. Deecoder for large vocabulary continuous speech recognition systems. Tehran: Sharif University of Technology; 2006. (Full Text in Persian)

21. Iran University of Science and Technology. Database analysis of speech and language speech recognition engines prepare feasibility. [Internet]. 2010 [cited 2012 Jan 12]. Available from: http://www.scict.ir/Portal/Home/Default.aspx?CategoryID=d7481bf6-47b2-456d-84ee-8e4651a8f5e1.

22. Mohr DN, Turner DW, Pond GR, Kamath JS, De Vos CB, Carpenter PC. Speech recognition as a transcription aid: a randomized comparison with standard transcription. J Am Med Informat Assoc 2003;10(1):85–93.

[Downloaded from pajoohande.sbmu.ac.ir on 2026-11-06]

23. Sheikhtaiefe M. A comparative study of speech recognition systems clinical data documenting the process in selected countries in 2010. Tehran: Tehran University of Medical Sciences; 2010. (Full Text in Persian)

24. Kirriemuir J. Speech recognition technologies. JISC, UK 2003;1(1):104–16.

25. Koerich AL, Sabourin R, Suen CY. Large vocabulary off-line handwriting recognition: a survey. Pattern analysis & applications. 2003;6(2):97–121.

26. Souto N, Meinedo H, Neto JP. Building language models for continuous speech recognition systems. Advances in Natural Language Processing: Springer; 2002. p. 101–10.

27. Shekofte Y, Almasghanj F, Merikhiahangarkalaee Y. Extraction of features including temporal patterns of speech signals using AR. Thirteenth National Conference of the Association Computer Iran; 2007 March 19-21; Kish Island, Persian Gulf, Iran. Sharif University of Technology, School of Computer Engineering; 2007.

28. Ramaswamy MR, Chaljub G, Esch O ,Fanning DD. Continuous speech recognition in MR imaging reporting advantages, disadvantages, and impact. Am J Roentgenol 2000;174(3):617–22.

29. Alapetite A. Impact of noise and other factors on speech recognition in anaesthesia. Int J Med Inform 2008;77(1):68–14.

30. Alapetite A. On speech recognition during anesthesia (dissertation). Denmark: Department of Communication, Business and Informations Technologies Roskilde; 2007.

31. Wagner S. Intralingual speech-to-text-conversion in real-time: challenges and opportunities. MuTra 2005; Challenges of multidimensional translation: Conference Proceedings; 2005.

32. Houston JD, Rupp FW. Experience with implementation of a radiology speech recognition system. J Digit Imaging 2000;13(3):124–8.

33. Molnar B, Gergely J, Toth G, Pronai L, Zagoni T, Papik K, *et al*. Development of a speech-based dialogue system for report dictation and machine control in the endoscopic laboratory. Endoscopy 2000;32(1):58–7.

34. Zafar A, Mamlin B, Perkins S, Belsito AM, Overhage JM, McDonald CJ. A simple error classification system for understanding sources of error in automatic speech recognition and human transcription. Int J Med Inform 2004;73(9):719–30.

35. Grasso MA. Structured speech input for clinical data collection. Proceedings of the 15th IEEE Symposium on Computer-Based Medical Systems (CBMS 2002); IEEE; 2002.

36. Jungk A, Thull B, Fehrle L, Hoeft A, Rau G. A case study in designing speech interaction with a patient monitor. J Clin Monit Comput 2000;16(4):295–307.

[Downloaded from pajoohande.sbmu.ac.ir on 2026-11-06]