

تعیین درجه اعتبار مفاصل و تخمین موقعیت مفاصل از دست رفته در داده‌های اسکلتی سنسور کینکت برای بهبود تشخیص ناهنجاری‌های اسکلتی

عطیه السادات میرمعینی^۱، حسن ختن‌لو^۲، وحید پورامین^۳، الهام علیقارداش^۴

چکیده

در سال‌های اخیر علاقه‌مندی به حسگرهای کینکت به دلیل قیمت پایین و کاربردهای گسترده در حال افزایش است. این ابزار می‌تواند با بهره‌گیری از دادگان اسکلتی و بدون استفاده از نشانگرها، وضعیت استقرار بدن انسان را برآورد کند. مسدود شدن بدن انسان با دیگر اشیا و حرکت سریع مقابل کینکت از معضلات اساسی برآورد موقعیت مفاصل است. در این پژوهش دو گام برای حل چالش موجود در نظر گرفته شده است. نخست، راهکاری بر مبنای مدل‌های اندازه‌گیری جهت تعیین درجه اعتبار مفاصل استخراج شده از حسگر کینکت ارائه شده که به عنوان ویژگی اثرگذار به همراه موقعیت مفاصل در کلاس‌بند پیشینه-حاشیه در نظر گرفته شده است. در گام دوم براساس میزان اعتبار هر یک از مفاصل، تصمیم‌گیری شده و مفاصل از دست رفته تشخیص داده می‌شوند. سپس با استفاده از الگوریتم‌های بخش‌بندی بدن انسان مبتنی بر شبکه یادگیری عمیق، اعتباربخشی مفاصل صورت می‌گیرد. نتایج نشان می‌دهد انتخاب ویژگی‌های مناسب در گام اول جهت مقایسه فریم‌های متوالی نسبت به روش‌های موجود، بهبود قابل توجهی در دقت کلاس‌بند دارد. همچنین گام دوم نیز با اعمال اعتبار بخشی به دادگان تاثیر زیادی بر افزایش دقت روش‌هایی دارد که از دادگان اسکلتی حسگر کینکت به عنوان ویژگی‌های ورودی بهره می‌برند.

کلیدواژه‌ها

حسگر کینکت، موقعیت مفاصل، دادگان اسکلتی، سنجش اعتبار، مفاصل از دست رفته

۱ مقدمه

یکی از اهداف اصلی محیط‌های هوشمند و سیستم‌های نظارتی، بهبود کیفیت زندگی انسان است. فهمیدن نیازها و مقاصد کاربر، پشتیبانی بهتر سیستم‌های هوشمند را به همراه خواهد داشت. از طرف دیگر با نظارت بر رفتارهای حرکتی کاربر، این سیستم‌ها می‌توانند وی را از موقعیت‌های خطرناک و حرکاتی که انجام آن‌ها منجر به آسیب رسیدن به کاربر می‌شود آگاه سازند. یک محیط هوشمند، با سیستم نظارتی خودکار برای ارزیابی وضعیت قرارگیری بدن، راه حلی بالقوه برای صرفه‌جویی در هزینه‌های بالای آسیب‌های محل کار و بیماری‌های مرتبط است. یکی از

این مقاله در مردادماه ۱۳۹۹ دریافت، در بهمن‌ماه بازنگری و پذیرفته شد.

^۱ آزمایشگاه هوش و بینایی ربات، گروه مهندسی کامپیوتر، دانشگاه بوعلی سینا
رایانامه: atiye.mirmoini@gmail.com

^۲ آزمایشگاه هوش و بینایی ربات، گروه مهندسی کامپیوتر، دانشگاه بوعلی سینا
رایانامه: khotanlou@basu.ac.ir

^۳ گروه مهندسی کامپیوتر، دانشگاه سید جمال‌الدین اسدآبادی
رایانامه: v.pouramin@sjau.ac.ir

^۴ گروه مهندسی کامپیوتر، دانشگاه سید جمال‌الدین اسدآبادی
رایانامه: e.alighardash@sjau.ac.ir

نویسنده مسئول: حسن ختن‌لو

چالش‌های مهم در اینگونه محیط‌ها، تحلیل فعالیت و وضعیت قرارگیری کاربر است.

از طرفی یکی از زمینه‌های تحقیقاتی مهم در جهت هوشمندسازی محیط، بررسی ناهنجاری‌های اسکلتی عضلانی است که درصد قابل توجهی از بیماری‌های مربوط به کار را به خود اختصاص می‌دهد. به دلیل اینکه دادگان اسکلتی به عنوان ورودی اصلی سیستم‌های مراقبت سلامت به شمار می‌آیند، در این پژوهش هدف بر آن است که با اعتبارسنجی و اعتباربخشی به دادگان اسکلتی استخراج شده از حسگر کینکت، ارزیابی وضعیت قرارگیری کاربران را از نظر سلامت جسمانی بهبود بخشد.

جمع‌آوری اطلاعات حرکتی، روشی مرسوم برای بدست آوردن وضعیت قرارگیری کاربر است. اطلاعات حرکتی انسان نقش اصلی را در سیستم‌های ارزیابی وضعیت قرارگیری (استقرار) بدن ایفا می‌کنند. در سال‌های اخیر دستگاه‌ها و سیستم‌های تجاری مختلفی به تکنولوژی برآورد وضعیت قرارگیری بدن انسان مجهز شده‌اند. این سیستم‌ها شامل حسگر کینکت، leap motion [۱]، body mounted camera [۲]، اسکنر لیزری سه بعدی [۳] و منبع نور مادون قرمز [۴] هستند که در اصول پیاده‌سازی و زمینه‌های کاربردی کاملاً متفاوت هستند [۵].

یکی از روش‌های جمع‌آوری اطلاعات مرتبط با حالات سه بعدی بدن استفاده از سیستم‌های ضبط حرکت پوشیدنی است که در بازی‌های کامپیوتری و انیمیشن‌ها مورد استفاده قرار می‌گیرد [۶]. این سیستم‌ها می‌توانند در یک محیط هوشمند اطلاعاتی را برای گزارش وضعیت‌های حساس بدن گزارش کنند [۷]. سیستم‌های پوشیدنی مختلف با قدرت و ضعف متفاوت وجود دارند که قادر هستند اطلاعات حرکتی و وضعیت قرارگیری سه بعدی بدن انسان را با استفاده از مجموعه‌ی نشانگرهایی که به بدن کاربر متصل می‌شوند، جمع‌آوری کنند [۸]. این روش در صورتی موفقیت‌آمیز است که نشانگرها برای دوربین قابل مشاهده باشد. سیستم‌های جمع‌آوری اطلاعات حرکتی مبتنی بر شتاب سنج [۹] تا [۱۰] و مبتنی بر آهنربا نیز با همین محدودیت روبرو هستند. با این وجود در این سیستم‌ها کاربر ملزم به پوشیدن حسگرها است و این امر باعث می‌شود که این تکنیک‌ها برای گسترش و استفاده در محیط‌های زندگی روزمره با دشواری همراه باشند.

روش دیگر تشخیص وضعیت قرارگیری بدن انسان استفاده از دوربین‌های ویدئویی معمولی و تجزیه و تحلیل تصاویر دو بعدی است [۱۱]. در این روش با ردیابی کردن بخش‌های غیر قابل تغییر در بدن انسان، وضعیت قرارگیری دو بعدی انسان در تصاویر ویدئویی تشخیص داده می‌شود [۱۲]. این تشخیص باعث می‌شود اطلاعات سطح بالایی مانند تعامل انسان و شی [۱۳] و هندسه‌ی صحنه نیز به دست آید [۱۴].

در سال‌های اخیر، دوربین‌های عمق سنج از جمله حسگرهای کینکت که مبتنی بر سیستم ردیابی حرکت هستند از محبوبیت بالایی برخوردار شده‌اند. این دوربین‌ها که انواع مختلفی دارند

[۱۵]، تصویر عمق را با استفاده از ساختار نور مادون قرمز بدست می‌آورند، سپس وضعیت قرارگیری بدن انسان می‌تواند برای مشخص کردن مفاصل مختلف بدن انسان ردیابی شود. این ردیابی با آموزش دادن یک درخت تصمیم که از پایگاه داده‌ی تصاویر عمق استفاده می‌کند صورت می‌پذیرد [۱۶] تا [۱۷]. دسته‌ی دیگری از روش‌های ردیابی استفاده از ابر نقاط حاصل از اسکنرهای سه بعدی هستند [۱۸] تا [۱۹]. با استفاده از دوربین‌هایی که تصویر عمق ایجاد می‌کنند، ردیابی بدون نیاز به تجهیزات پوشیدنی انجام می‌شود [۲۰].

در مقایسه با سیستم‌های ردیابی حرکت مبتنی بر اپتوالکترونیک یا الکترومغناطیس، حسگرهای کینکت بسیار کم هزینه، قابل حمل و بدون نیاز به اتصال نشانگر هستند. از طرفی حسگر کینکت به دلیل پشتیبانی از دادگان سه بعدی و قیمت ارزان مورد توجه بسیاری از محققان قرار گرفته است، با این وجود این حسگر در استخراج اطلاعات مربوط به دادگان اسکلتی که در پژوهش‌های مرتبط با محیط هوشمند بسیار مورد استفاده قرار می‌گیرد با چالش‌هایی مواجه است [۲۱] که سازوکار الگوریتم‌های یادگیری را با مشکل رو به رو می‌سازد.

در این پژوهش چهارچوبی برای کلاس‌بندی وضعیت‌های قرارگیری افراد بر اساس دادگان اسکلتی ارائه شده است. برای مقابله با دادگان اسکلتی دارای خطا به عنوان ورودی، مجموعه‌ای از متغیرهای اندازه‌گیری اعتبار مفاصل معرفی می‌شوند. این متغیرها میزان اعتبار هر یک از مفاصل ردیابی شده را ارزیابی می‌کنند، سپس براساس اعتبارسنجی انجام شده مفاصلی که از اعتبار کافی برخوردار نیستند به عنوان مفاصل از دست‌رفته شناخته می‌شوند.

به دلیل چالش موجود در مسئله، تاکنون روشی یکپارچه جهت تشخیص و تخمین مفاصل از دست‌رفته در حسگر کینکت ارائه نشده است و روش‌های ارائه شده با توجه به نوع کاربرد و حوزه‌ی استفاده‌ی آنها تعیین می‌شوند. بنابراین مطالعات انجام شده در سه حوزه‌ی اصلیت تشخیص مفاصل، تخمین مفاصل و بررسی ناهنجاری‌های اسکلتی عضلانی مطرح می‌شوند.

در این پژوهش به دستاوردهای زیر پرداخته شده است:

- تشخیص و مطالعه چالش‌های تصویربرداری اسکلتی به کمک حسگر کینکت
 - ارائه راهکار بر مبنای مدل‌های اندازه‌گیری جهت تعیین درجه‌ی اعتبار مفاصل استخراج شده از حسگر کینکت در جهت حل چالش‌ها
 - تشخیص میزان اعتبار هر یک از مفاصل و تصمیم‌گیری درباره مفاصل از دست‌رفته با استفاده از اطلاعات مرحله قبل و تشخیص ناهنجاری‌های اسکلتی بدن
- در ادامه در بخش ۲ مطالعات پیشین بررسی خواهد شد. سپس در بخش ۳ روش پیشنهادی بیان می‌شود که در بخش ۴ مورد تحلیل

از اندازه‌گیری‌ها را ارائه می‌دهند که به طور عینی درجه‌ی اعتبار هر بخش ردیابی شده از بدن را ارزیابی می‌کند. با ترکیب درجات اعتبار بدست آمده با پایگاه داده‌ی مربوطه، مجموعه‌ای از حالت‌های مشابه که از نظر حرکتی معتبر هستند بدست می‌آید.

۲-۲ تخمین مفاصل بدن انسان

در پژوهشی که در سال ۲۰۱۱ ارائه شده، یک تکنیک مهم و مستقل از دید برای تخمین مفاصل مورد بررسی قرار گرفته است [۱۸]. در این مطالعه با استفاده از یک الگوریتم منطبق سازی، شبکه‌ای سه بعدی از سطح کامل بدن را بر روی تصویر عمق نگاشت می‌کنند. در این روش از یک پایگاه‌داده‌ی حرکتی استفاده شده است که توسط یک تولید کننده‌ی مدل شبکه‌ای انسان بوجود آمده است.

شکل ۱ طرح کلی از روند پردازشی این مطالعه را نشان می‌دهد. ابتدا ابر نقاط داده شده که حاصل تصویربرداری عمقی با حسگر کینکت است، تشکیل می‌شود. سپس اشیای نامرتب بر اساس اطلاعات فاصله‌ای حذف می‌شوند. در گام بعد، الگوریتم بازسازی و اصلاح سطح جهت حذف نویزها اعمال می‌شود. ابر نقاط پالایش شده به یک قالب هماهنگ و متعارف جهت حذف وابستگی زاویه دید انتقال داده می‌شود. در مرحله بعد یک حالت اسکلتی، نزدیک به تصویر حاصل از مجموعه پایگاه داده حرکتی تولید شده، انتخاب می‌شود. در نهایت پیکربندی مناسبی بین حالت انتخابی و تصویر عمق ارائه شده انجام شده و موقعیت مفاصل از نگاشت ساختار اسکلتی بر روی تصویر عمق تخمین زده می‌شوند.

در مطالعه شفافی و همکارانش در سال ۲۰۱۶ [۲۵] روشی برای برآورد حالت ارائه شده است. آن‌ها با مشخص کردن محل بخش‌های مختلف بدن، بدون نیاز به مدل صریح شکل‌ها، موقعیت مفاصل را با ترکیب مستندات که از چندین زاویه‌ی دید بدست آمده در زمان واقعی مشخص می‌کنند. برای این کار با استفاده از پارامترهای ذاتی هر کینکت، ابر نقاط محلی بازسازیشده، سپس برای انتقال هر نقطه‌ی ابری به یک فضای مرجع، تمام نقاط ابری با استفاده از پارامترهای ذاتی دوربین ادغام می‌شوند.

و بررسی قرار خواهد گرفت. در نهایت در بخش ۵ به نتیجه‌گیری و جمع‌بندی پرداخته خواهد شد.

۲ کارهای پیشین

چنانچه ذکر شد به دلیل عدم ارائه روشی یکپارچه جهت تشخیص و تخمین مفاصل از دست رفته در حسگر کینکت، مطالعات انجام شده در سه حوزه‌ی اصلی مطرح می‌شوند که در ادامه به آن‌ها پرداخته می‌شود.

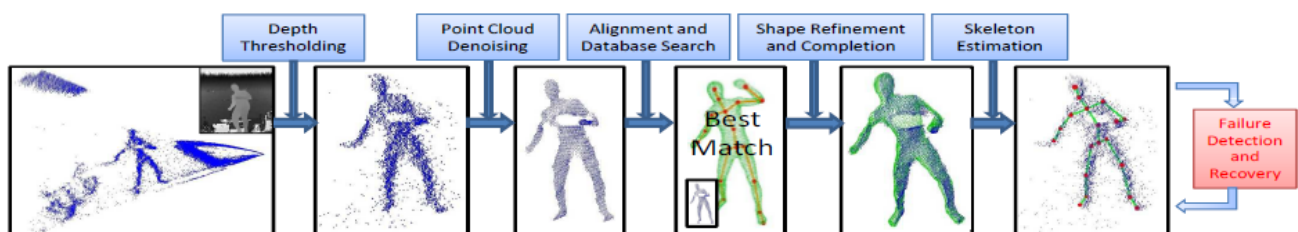
۱-۲ اعتبار سنجی دادگان اسکلتی استخراج شده از حسگر کینکت

میزان قابلیت اطمینان و دقت دستگاه‌های RGB-D از جمله کینکت در چندین مطالعه مورد بررسی قرار گرفته است [۲۲] تا [۲۳]. این مطالعات نشان می‌دهند استخراج دادگان اسکلتی از اطلاعات تصویر عمق به دلیل چالش‌هایی از جمله عدم دقت و مسدود شدن بخش‌هایی از بدن با اشیای دیگر موجود در صحنه، صحت خود را از دست می‌دهد. همچنین اطلاعات بدست آمده از حسگر کینکت به شدت وابسته به زاویه‌ی فرارگیری حسگر است. از این رو اعتبارسنجی دادگان اسکلتی به دو روش مورد بررسی قرار می‌گیرد:

- استفاده از نشانگرهای متصل به بدن
- متغیرهای سنجش اعتبار

در پژوهش ایدرزالک و همکاران [۲۲] به طور همزمان اطلاعات حرکتی با استفاده از یک سیستم ضبط اطلاعات مبتنی بر نشانگر و حسگر کینکت جمع‌آوری شده است. در این روش از ۴۳ نشانگر استفاده می‌شود که تقریباً در نواحی مشخص و استاندارد از بدن قرار می‌گیرند. در این مطالعه با استفاده از نرم افزارهای تجاری Phase space recap و Autodesk motion builder از دو روش مختلف جهت ایجاد ساختار اسکلتی استفاده شده است که در آن ردیابی بدن در کاربردهای مرتبط با نظارت بر سلامت ممکن است. اما یکی از مشکلات اصلی حسگر کینکت عدم ارائه‌ی اطلاعات صحیح به هنگام مسدود شدن اندام‌ها است که باید با الگوریتم‌های مناسب کنترل شود.

از جمله روش‌های مبتنی بر متغیرهای سنجش اعتبار، مطالعه شوم و همکاران است [۲۴]، که در آن روشی جدید برای بازسازی حرکات متغیر از روی حالت‌های ناقص و دارای نویز که توسط کینکت جمع‌آوری شده است ارائه داده‌اند. آن‌ها ابتدا مجموعه‌ای



شکل ۱ طرح کلی روش ارائه شده در مطالعه [۱۸]

دادگان اسکلتی حسگر کینکت در افراد براساس پایگاه داده محیط کار پرداخته می‌شود. شکل ۲ بلوک دیاگرام روش پیشنهادی را نشان می‌دهد.

در روش پیشنهادی از کلاس‌بند با نظارت بهره‌گرفته، و برچسب‌گذاری نمونه‌های موجود در پایگاه داده‌ی مورد نظر براساس قوانین ارائه شده از سوی آژانس اروپا برای سلامت و ایمنی کار انجام شده است [۲۹].

۳-۱ متغیرهای سنجش اعتبار

میزان اعتبار داده‌های منبع می‌تواند توسط مجموعه‌ای از روش‌های اکتشافی اندازه‌گیری شود. در ادامه روش‌های پیشنهادی جهت اعتبار سنجی مفاصل مطرح می‌گردد.

۳-۱-۱-۱ متغیر سنجش اعتبار مفاصل براساس فریم رنگی

متغیر سنجش اعتبار براساس فریم‌های رنگی، میزان اعتبار هر مفصل را براساس نزدیکی ویژگی‌های استخراج شده بین دو فریم همسایه در یک ویدئو رنگی ارزیابی می‌کند. به دلیل اینکه حرکات انسان به طور طبیعی پیوسته است، ظاهر مفاصل در فریم‌های رنگی همسایه باید از لحاظ بصری مشابه باشد. معمولاً ظاهر متفاوت مفاصل در میان فریم‌های مجاور بیانگر این است که مفصل مورد نظر در حداقل یک فریم به اشتباه ردیابی شده است. در روش پیشنهادی میزان اعتبار یک مفصل توسط استخراج ناحیه‌ای از پیکسل‌های تصویر رنگی به مرکزیت مفصل مورد بررسی محاسبه می‌شود، سپس این نواحی در فریم‌های مجاور توسط ویژگی‌های استخراج شده از آن‌ها مورد ارزیابی قرار می‌گیرد.

روش‌های بسیاری برای استخراج ویژگی و مقایسه‌ی دو ناحیه‌ی استخراج شده از دو فریم مجاور یک ویدئو وجود دارد، اما بهترین نمایش ویژگی و معیار مقایسه قطعاً به دامنه‌ی کاربردی مسئله بستگی دارد. در روش پیشنهادی از دو ویژگی هیستوگرام جهت‌گرادیان^۱ و همبستگی رنگی^۲ استفاده می‌شود که از مقاله [۲۸] ایده گرفته شده است.

- اختلاف هیستوگرام جهت‌گرادیان نواحی استخراج شده از مفصل مورد نظر در فریم‌های مجاور

ورودی توصیف‌گر هیستوگرام جهت‌گرادیان ناحیه‌ای با سایز $3 \times 50 \times 50$ به مرکزیت مفصل مورد نظر است. برای استخراج ویژگی، هر ناحیه به سلول‌هایی با سایز 8×8 پیکسل تقسیم می‌شود، سپس هر گروه 2×2 از سلول‌ها به عنوان یک بلاک در نظر گرفته می‌شود.

در مطالعه فرنوش و همکاران [۲۶] روشی برای بازسازی مفاصل از دست‌رفته‌ها استفاده‌از یک تصویر عمق ارائه شده است. در این روش از ترکیب نزدیکترین مفصل به مفصل از دست‌رفته‌ها برچسب‌های روی بدن استفاده شده است.

۳-۲ ارزیابی وضعیت قرارگیری بدن توسط حسگر کینکت و بررسی ناهنجاری اسکلتی عضلانی

بررسی ناهنجاری‌های اسکلتی عضلانی، به ارزیابی وضعیت قرارگیری بدن که در معرض عوامل خطرناک مربوط به محیط کار می‌باشد، می‌پردازد.

لیو و همکارانش [۲۷] روشی برای برآورد وضعیت قرارگیری مناسب کاربر در حالت نشسته در صحنه‌های RGB-D با استفاده از زمینه‌ی اشیا ارائه داده‌اند. این روش با بهره‌گیری از روابط فضایی و زمینه‌ای میان اشیا و مفاصل به بررسی ناهنجاری‌های اسکلتی-عضلانی می‌پردازد. برای انجام این کار، وضعیت قرارگیری افراد در حالت نشسته در یک محیط اداری در تعامل با رایانه ارزیابی می‌شود. در گام اول یک پیکربندی سالم برای حالت نشستن کاربر بر روی صندلی از دیدگاه ارگونومی تعریف می‌شود. در گام دوم روابط فضایی و زمینه‌ای بین اشیا و مفاصل در صحنه RGB-D یاد گرفته می‌شود. در گام سوم از کلاس‌بند NaiveBayes برای کلاس‌بندی وضعیت‌های سالم و ناسالم استفاده می‌شود. یک الگوریتم تکراری وزن‌دار برای محاسبه‌ی مرکز ثقل (COF) مفاصلی که در ساختار اسکلتی سالم کلاس‌بندی شده‌اند استفاده می‌شود. در گام آخر مرکز ثقل‌های محاسبه شده به عنوان موقعیت‌های نهایی هر مفصل در نظر گرفته می‌شوند. در نهایت از طریق پالایش توزیع مفاصل، ساختار اسکلتی بازسازی می‌شود.

هو و همکارانش [۲۸] یک چهارچوب جدید برای نظارت و کلاس‌بندی وضعیت قرارگیری بدن کاربر پیشنهاد داده‌اند. ابتدا میزان اعتبار مفاصل مشاهده شده از کینکت توسط متغیرهای اعتبارسنجی حرکتی و رفتاری ارزیابی می‌شود. سپس مقادیر اعتبار بدست آمده برای هر یک از مفاصل به عنوان وزن در یک customized max-margin برای کلاس‌بندی دقیق داده‌ها با نوبز بالا به کار گرفته می‌شود.

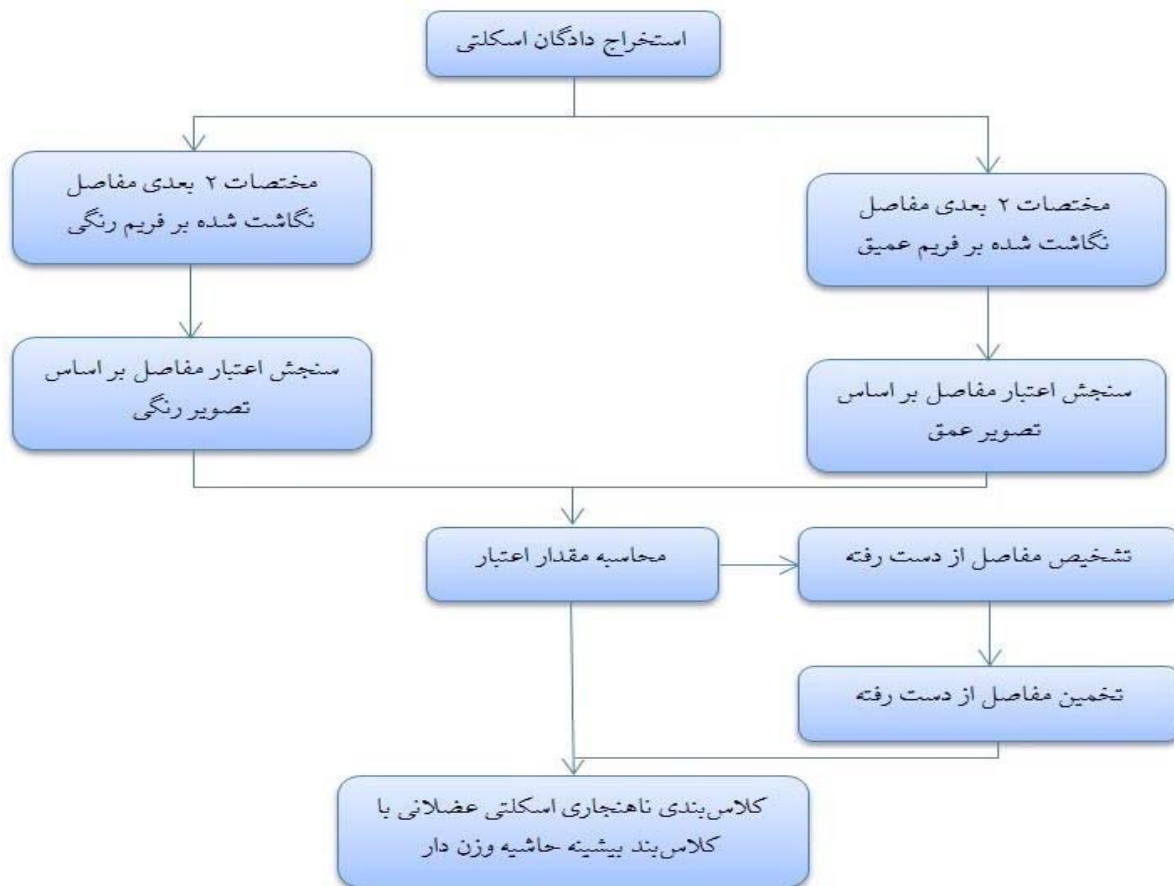
۳ روش پیشنهادی

روش پیشنهادی در این مقاله به منظور تشخیص و تخمین مفاصل از دست رفته و تعیین درجه اعتبار مفاصل براساس دادگان اسکلتی استخراج شده از حسگر کینکت است. در این روش هدف، بهبود دقت کلاس‌بندی‌هایی است که دادگان اسکلتی را به عنوان ورودی می‌پذیرد. در ادامه به بررسی روش پیشنهاد شده برای تشخیص و تخمین مفاصل از دست رفته و تعیین درجه اعتبار مفاصل در

¹Heuristics

²Histogram of Oriented Gradient

³Color Correlogram



شکل ۲ بلوک دیاگرام روش پیشنهادی

است [۳۰]. در روش پیشنهادی با بهره‌گیری از این ایده و استفاده از ویژگی همبستگی رنگی به مقایسه‌ی نواحی استخراج شده از فریم‌های مجاور پرداخته می‌شود.

از آنجایی که هیستوگرام‌های رنگی اطلاعات فضایی را در بر نمی‌گیرند نتایج ضعیفی را در زمینه‌ی بازیابی و مقایسه‌ی تصاویر ارائه می‌دهند. در ویژگی همبستگی رنگی، اطلاعات رنگی و فضایی تصویر با یکدیگر ترکیب می‌شوند، در حالی که مزایای هیستوگرام‌ها نیز حفظ می‌شود. در این بخش برای محاسبه‌ی میزان اعتبار هریک از مفاصل از شباهت کسینوسی بردار همبستگی رنگی نواحی استفاده می‌شود. رابطه (۲) نحوه‌ی محاسبه‌ی میزان اعتبار را بیان می‌کند. به طوری که Rc_i برای هر دو توصیف‌گر مقداری بین بازه [۰ و ۱] دارد. ورودی الگوریتم سنجش اعتبار مفاصل فریم‌هایی از دادگان ۲ بعدی مفاصل و فریم‌های رنگی مربوطه و خروجی آن اعتبار هریک از مفاصل در بازه [۰ و ۱] است [۲۸].

$$Rc_i = \frac{\text{Corr-vector } i.f \bullet \text{Corr-vector } i.f+1}{\|\text{Corr-vector } i.f\| \|\text{Corr-vector } i.f+1\|} \quad (2)$$

در رابطه (۲)، $\text{Corr-vector } i.f$ معادل بردار همبستگی رنگی مفصل i در فریم f است. به همین ترتیب $\text{Corr-vector } i.f+1$ معادل بردار همبستگی رنگی مفصل i در فریم $f+1$ است. در مخرج

جهت جلوگیری از اثر تغییرات جزئی رنگی برای هریک از نواحی در فریم‌های مجاور هیستوگرام جهت‌گردان هشت بین محاسبه می‌شود. در نهایت میزان اعتبار براساس فریم رنگی برای هر مفصل i با استفاده از فرمول شباهت کسینوسی بین دو بردار ویژگی استخراج شده از دو ناحیه در فریم‌های مجاور براساس رابطه (۱) محاسبه می‌شود [۲۸].

$$Rc_i = \frac{\text{Hog-vector } i.f \bullet \text{Hog-vector } i.f+1}{\|\text{Hog-vector } i.f\| \|\text{Hog-vector } i.f+1\|} \quad (1)$$

در رابطه (۱)، $\text{Hog-vector } i.f$ معادل بردار هیستوگرام جهت‌گردان مفصل i در فریم f است. به همین ترتیب $\text{Hog-vector } i.f+1$ معادل بردار هیستوگرام جهت‌گردان مفصل i در فریم $f+1$ است. در مخرج رابطه نیز از اندازه این بردارها استفاده شده است.

- اختلاف همبستگی رنگی نواحی استخراج شده از مفصل مورد نظر در فریم‌های مجاور

یکی از روش‌های قدرتمندی که برای بازیابی تصاویر و ویدئو ارائه شده است استفاده از همبستگی رنگی به جای هیستوگرام

سازد [۳۱]. همچنین اطلاعات نادرست می‌تواند مانع از کلاس‌بندی عناصر صحیح نیز شود [۳۲]. بنابراین با توجه به اهمیت صحت عناصر بردار ویژگی، در این پژوهش فرآیند اعتباربخشی بر روی دادگان اسکلتی اعمال می‌شود و تاثیر آن بر نتیجه‌ی کلاس‌بندی مورد بررسی قرار می‌گیرد.

۳-۲-۱ تشخیص مفاصل از دست‌رفته

به ازای هر دو حالت مورد بررسی (نشسته و ایستاده) یک ماتریس اعتبار، با توجه به متغیر سنجش اعتبار مفاصل براساس فریم عمق، به صورت جدا محاسبه می‌شود. سپس عناصر ماتریس مورد ارزیابی قرار گرفته و چنانچه مقدار یک عنصر از حد آستانه کمتر باشد به عنوان مفصل از دست‌رفته برچسب زده می‌شود. در ادامه، مفاصل با توجه به ماتریس‌های اعتبار بدست آمده از مرحله اعتبارسنجی فیلتر شده و مفاصل با اعتبار پایین با استفاده از الگوریتم بخش‌بندی بدن که در ادامه بیان خواهد شد تخمین زده می‌شوند.

۳-۲-۲ بخش‌بندی بدن انسان

در این بخش از روش پیشنهادی با الهام از مقاله از یک شبکه CNN برای کلاس‌بندی تراکم در تصاویر عمق بهره‌گرفته می‌شود. معماری شبکه که از پژوهش لانگ و همکارانش [۳۳] ایده‌گرفته شده است در شکل ۳ مشاهده می‌شود.

خروجی شبکه ۴۴ ناحیه شامل ۴۳ ناحیه مربوط به بخش‌های مختلف بدن و ۱ ناحیه مربوط به پس‌زمینه است. این نواحی و نحوه‌ی بخش‌بندی آن‌ها در شکل ۴ نشان داده شده است.

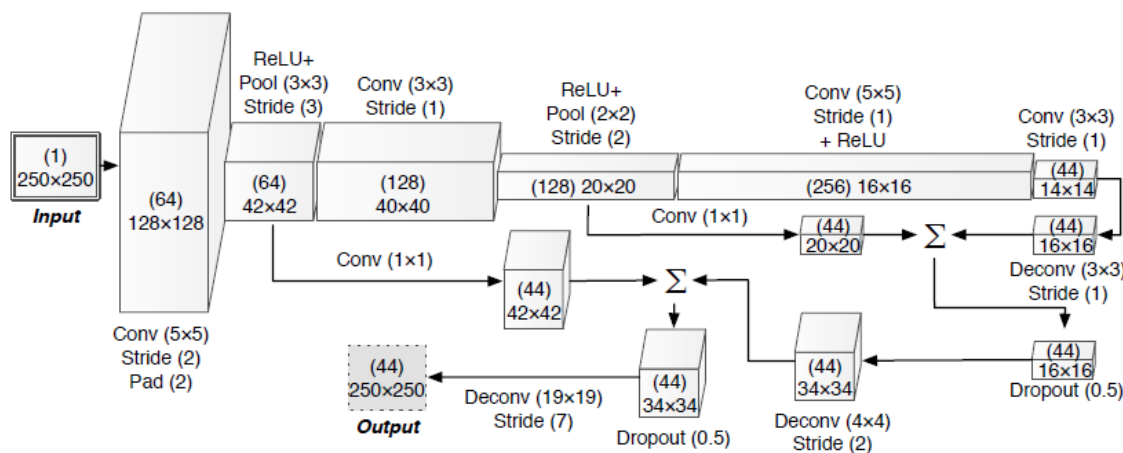
رابطه نیز از اندازه این بردارها استفاده شده است. یک همبستگی رنگی بیان می‌کند که همبستگی فضایی بین جفت رنگ‌ها با چه فاصله‌ای تغییر می‌کند. به طور غیر رسمی همبستگی برای یک تصویر جدولی است که با جفت‌های رنگی نشان داده می‌شود. به طوری که d آمین ورودی برای سطر (i, j) احتمال یافتن پیکسل رنگ z در فاصله d از پیکسل رنگ i را در این تصویر مشخص می‌کند و d از مجموعه مقادیر مفاصل انتخاب شده است.

۳-۱-۲ متغیر سنجش اعتبار مفاصل براساس فریم عمق

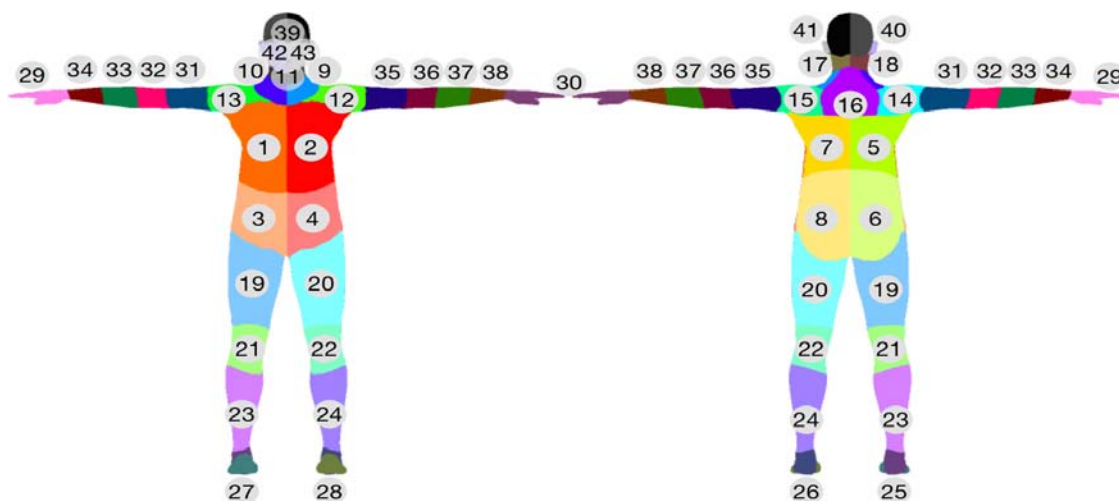
سنجش براساس فریم عمق، میزان اعتبار هر یک از مفاصل را بر اساس نزدیکی ویژگی استخراج شده بین دو فریم مجاور در یک توالی از تصویر عمق ارزیابی می‌کند. ایده‌ی متغیر اعتبارسنجی براساس فریم عمق این است که تغییرات ناگهانی در میان دو فریم مجاور عمق بر روی موقعیت تشخیص داده شده‌ی مفاصل اثر می‌گذارد و نشان دهنده‌ی ردیابی نادرست مفصل در ناحیه‌ی دارای تغییرات زیاد است. در این بخش نیز هیستوگرام جهت‌گرادیان هشت بین نواحی محاسبه می‌شود و براساس شباهت کسینوسی بردارهای ویژگی میزان اعتبار مفاصل سنجیده می‌شود (فرمول (۱)). در این بخش نیز همانند استفاده از فریم رنگی نواحی به مرکزیت مفاصل عمل شده ولی از فریم عمق استفاده می‌شود.

۳-۲ اعتباربخشی به مفاصل

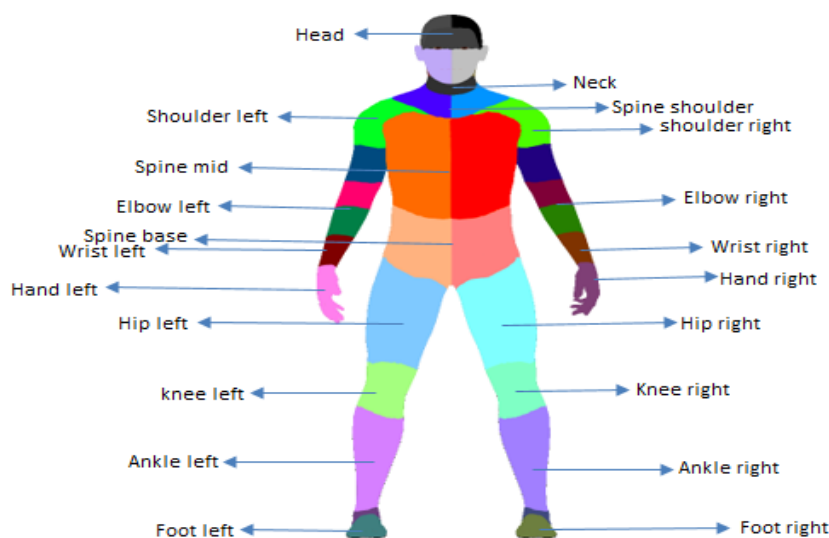
از آنجایی که بردارهای ویژگی اساس یک کلاس‌بندی را تشکیل می‌دهند، اطلاعات صحیح می‌تواند کلاس‌بندی‌هایی که بر پایه الگوریتم‌های یادگیری ماشین هستند را به سمت نرخ موفقیت بالا سوق دهد. از طرف دیگر بردار ویژگی شامل اطلاعات غیرمرتبط و نادرست می‌تواند سازوکار الگوریتم‌های کلاس‌بندی را پیچیده



شکل ۳ معماری شبکه CNN کلاس‌بند تراکم [۳۳]



شکل ۴ نواحی مربوط به بخش بندی بدن انسان



شکل ۵ تخمین مفاصل براساس خروجی شبکه CNN

۳-۳ تخمین مفاصل

تخمین مفاصل طی مراحل زیر انجام می‌گیرد: در گام اول تصاویر عمق و شاخص بدن را از حسگر Kinect جمع آوری کرده، پس از نگاهت شاخص بدن بر تصویر عمق، محدوده‌ی بدن انسان از تصویر عمق استخراج می‌شود. در گام دوم تصویر خروجی از گام اول به عنوان ورودی به شبکه‌ی CNN اشاره شده در شکل ۳ داده می‌شود. در گام سوم برای تخمین موقعیت مفاصل با توجه به ساختار اسکلتی بدن انسان، متناسب با هر یک از مفاصل میانگین موقعیت پیکسل‌ها را در کلاس معین در نظر می‌گیریم. شکل ۵ تخصیص ساختار اسکلتی براساس خروجی شبکه را نشان می‌دهد.

۳-۴ کلاس بندی ناهنجاری اسکلتی عضلانی

در بسیاری از مسائل کلاس بندی، ویژگی‌های مفید تحت تاثیر ویژگی‌های کم اهمیت و با تعداد بیشتر قرار می‌گیرند و موجب می‌شوند کارایی الگوریتم کلاس بندی به شدت کاهش یابد. ماشین بردار پشتیبان معمولی هنگامی که به طور مستقیم برای اینگونه مسائل اعمال می‌شود عملکرد ضعیفی را نشان می‌دهد. از آنجایی که دادگان مورد بررسی در این پژوهش دادگان اسکلتی بوده و هر یک از درجه‌ی اعتبار متفاوتی برخوردار هستند، سعی شده تا مفاصل با میزان اعتبار بالاتر، به منظور بهبود نتایج کلاس بندی، نقش بیشتری در الگوریتم ایفا کنند. بنابراین برای کلاس بندی ناهنجاری‌ها که براساس دادگان اسکلتی استخراج شده از حسگر کینکت می‌باشند از یک ماشین بردار پشتیبان وزن دار (Weighted SVM) استفاده می‌شود. در این بخش از پارامترها و متغیرهای مطالعه [۲۸] استفاده شده است.

یک ماشین بردار پشتیبان کلاسیک براساس رابطه (۳) تعریف می‌شود [۲۸]:

$$\rho(i) = y_i \cdot w \frac{R_i}{\|R_i\|} x_i \quad (7)$$

به طوری که بردار وزن R_i برای نمونه i توسط $\|R_i\|$ نرمال می‌شود. در نتیجه ویژگی‌هایی مقدار اعتبار بالاتر در محاسبه‌ی حاشیه‌ی نمونه بیشتر شرکت می‌کنند.

۴ تحلیل و بررسی

۴-۱ پیاده سازی

در بخش متغیرهای سنجش اعتبار برای محاسبه‌ی اعتبار رفتاری مفاصل از مختصات سه بعدی دادگان اسکلتی استفاده می‌شود. محاسبه‌ی اعتبار هر مفصل در نواحی با سایز $50 * 50$ پیکسل به مرکزیت مفصل در فریم‌های رنگی و عمق انجام می‌گیرد.

برای بخش بندی بدن انسان از شبکه CNN با ورودی تصاویر عمقی استفاده شده است. برای پیاده سازی این شبکه از ابزار Tensor Flow در بستر Windows 10 بهره گرفته شده است. آموزش و تست این شبکه نیز در یک کامپیوتر با پردازنده Core i7، حافظه اصلی ۱۶ GB و پردازنده گرافیکی NVIDIA 1080Ti انجام شده است. این شبکه توسط دادگان مصنوعی UBC3V [۲۵] آموزش داده می‌شود. UBC3V شامل تصاویر عمق گرفته شده از کینکت است که برای آموزش شبکه‌های تخمین مفاصل استفاده می‌شود.

ابتدا تلاش شده است که مستقیماً شبکه با مجموعه داده‌ی Hard-Pose (جدول ۱) آموزش داده شود، با این حال در تمام آزمایشات با بهینه سازی‌های مختلف دقت شبکه به طور متوسط در کلاس بندی هر کلاس بیشتر از ۵۰ درصد نمی‌شود. بر اساس روش بنژیو و همکارانش [۳۴] روند آموزش با معرفی مجموعه دادگان ساده Easy-Pose و Inter-Pose (جدول ۱) که دارای درجه پیچیدگی متوسط است، انجام می‌شود. بنابراین آموزش شبکه با Easy-Pose آغاز شده و در هر تکرار ۸ تصویر عمق پردازش می‌شود، سپس تنظیمات دقیق پارامترهای مربوط به نتایج شبکه‌ی اولیه بر روی مجموعه دادگان Inter-Pose اعمال می‌شود. برای بخش بندی بدن انسان از ابزار Matconvnet beta 25_1.0 استفاده شده است.

۴-۲ پایگاه داده‌ها

سه مجموعه داده با حالات اسکلتی ساده، متوسط و دشوار برای آموزش شبکه تعریف می‌شود. ویژگی این سه مجموعه در جدول ۱ آمده است.

$$\begin{aligned} \text{Min}_{w, \xi, b} & \left(\frac{1}{2} \|W\|^2 + C \sum_{i=1}^n \xi_i \right) \\ \text{subject to} & y_i(w x_i + b) \geq 1 - \xi_i \quad i = 1 \dots n \end{aligned} \quad (3)$$

به طوری که x_i و y_i به ترتیب ویژگی‌ها و برچسب نمونه هستند. C پارامتری برای معامله^۱ میان پیچیدگی مدل و دقت است، b یک آستانه و ξ متغیری است برای مدیریت نمونه‌های آموزشی که به طور خطی غیرقابل جدا شدن هستند. به طور خاص، w با به حداکثر رساندن حاشیه ρ طبق رابطه (۴) یادگرفته می‌شود [۲۸].

$$\rho = \min_i y_i (W \cdot X_i + b) / \|W\|^2 \quad (4)$$

با این حال هنگامی که نمونه‌هایی با ویژگی از دست رفته مدیریت می‌شود کل بردار ویژگی x_i در محاسبه‌ی حاشیه برای آموزش کلاس بند شرکت داده می‌شود، در واقع فرآیند آموزش بدون نادیده گرفتن ویژگی‌های غائب انجام می‌شود (معمولاً ویژگی از دست رفته با یک مقدار پیش بینی شده و یا به سادگی با صفر جایگزین می‌شود)، در نتیجه عملکرد کلاس بندی یادگیری شده تضعیف می‌شود.

برای کلاس بندی دادگان با ویژگی‌های غائب، هر نمونه با محاسبه‌ی حاشیه‌ی نمونه در زیرمجموعه‌ای از فضای ویژگی‌های خودش نسبت به فضای کلی ویژگی که در رابطه (۵) بیان شده است، در نظر گرفته می‌شود [۲۸]:

$$\rho(i) = \frac{y_i \cdot w(i) \cdot x_i}{\|w(i)\|} \quad (5)$$

که $w(i)$ شامل زیرمجموعه‌ای از کلیه‌ی مقادیر w مرتبط با ویژگی‌های معتبر در x_i است، در این حالت حاشیه کلاس بند با به حداقل رساندن حاشیه نمونه‌ها طبق رابطه (۶) است [۲۸]:

$$\text{Max}_w \left(\min_i \frac{y_i \cdot w(i) \cdot x_i}{\|w(i)\|} \right) \quad (6)$$

در روش پیشنهادی بردار وزن R_i دارای ابعاد یکسان با بردار ویژگی برای یک نمونه i (یک وضعیت قرارگیری) است. R_i ، z ، R_i وزن یک ویژگی اسکلتی z (مفصل z) است. در این بخش بردار وزنی برای دو متغیر سنجش اعتبار مفاصل بر اساس فریم رنگی و عمق در نظر گرفته می‌شود، بنابراین حاشیه نمونه براساس رابطه (۷) محاسبه می‌شود [۲۸]:

¹Tradeoff

جدول ۱ ویژگی مجموعه دادگان آموزشی مصنوعی

پایگاه داده	تعداد حالات	تعداد نفرات	تعداد نمونه
Easy_Pose	۱۰ هزار	۱	1 میلیون
Inter_Pose	۱۰۰ هزار	۱	1.3 میلیون
Hard_Pose	۱۰۰ هزار	۱۶	300 هزار

جدول ۲ نتایج کلاس‌بند وزن‌دار روش پیشنهادی براساس فریم عمق

دقت	حساسیت	اختصاصیت	میانگین هندسی
۹۸/۴۰	۹۶/۸۵	۹۹/۰۰	۹۷/۹۸
۹۷/۱۸	۹۷/۸۶	۹۶/۷۲	۹۷/۲۷

جدول ۳ مقایسه‌ای بین روش پیشنهادی با سایر روش‌های پیاده‌سازی شده بر پایگاه داده health را در وضعیت قرارگیری ایستاده و نشسته نشان می‌دهد. روش پایه برای کلاس‌بندی وضعیت‌های قرارگیری میزان اعتبار مفاصل را در نظر نمی‌گیرد و تنها براساس موقعیت مفاصل قضاوت می‌کند. در روش دیگر که وانگ و همکارانش [۳۵] ارائه داده‌اند از بردار ویژگی متشکل از موقعیت مفاصل و موقعیت‌های نسبی بین هر جفت مفصل استفاده می‌کند. روش دیگر که مورد مقایسه قرار می‌گیرد روش هو و همکارانش [۲۸] است که در این روش برای محاسبه‌ی متغیرهای اعتبار رنگی براساس فریم رنگی و عمق از گرادیان پیکسل‌های تصویر استفاده شده است.

جدول ۳ مقایسه روش پیشنهادی در وضعیت‌های نشسته و ایستاده

روش	میانگین دقت	میانگین دقت
	حالت ایستاده	حالت نشسته
Joint Position	۸۰/۸۴	۶۶/۶۷
روش پایه هو [۲۸]	۸۶/۳۲	۷۰/۵۸
راهکار پیشنهادی هو [۲۸]	۸۸/۶۷	۷۹/۴۵
سنجش اعتبار مفاصل در تصاویر رنگی با استفاده هیستوگرام گرادیان	۹۹/۰۲	۹۸/۵۵
سنجش اعتبار مفاصل در تصاویر رنگی با Correlogram	۹۸/۶۸	۹۶/۴۷
سنجش اعتبار مفاصل در تصاویر عمق با استفاده از هیستوگرام گرادیان	۹۸/۴۰	۹۷/۱۸

همانطور که در جدول نتایج ۳ مشاهده می‌شود کمترین میانگین دقت از روشی که تنها بر پایه‌ی موقعیت مفاصل است، به دست می‌آید. در این صورت دادگان نامعتبر نیز نقشی برابر با دادگان معتبر ایفا می‌کنند که همین امر موجب پیچیده شدن ساز و کار کلاس‌بند می‌شود. در روش ادموند که ترکیبی از روش وانگ به همراه سنجش اعتبار مفاصل است اندازه‌گیری براساس گرادیان پیکسل‌های موجود در ناحیه‌ای به سایز 27×27 به مرکزیت مفصل می‌باشد. بردار گرادیان حاصل، تغییرات مقدار پیکسل‌ها از یک پیکسل نسبت به پیکسل‌های اطراف را نشان می‌دهد، که به شدت به تغییرات مقدار پیکسل‌ها حساس است و میزان اعتبار مفاصل را تحت تاثیر قرار داده و مفاصل معتبر را کم اهمیت تلقی می‌کند، اما در Hog تعداد وقوع جهت‌گیری‌های گرادیان در یک قسمت

ساده‌ترین مجموعه (Easy_Pose) دارای یک شخص است که با زیر مجموعه‌ای از وضعیت‌های "قدم زدن" و "دویدن" برچسب‌گذاری شده است. مجموعه‌ی دوم (Inter_Pose) با افزایش پیچیدگی توسط وضعیت‌های متنوع، نیاز کلاس‌بند به مدیریت تغییرات زیاد از وضعیت‌های قرارگیری را برطرف می‌سازد، و در نهایت آخرین مجموعه داده (Hard_Pose) شامل تمام نفرات با شکل‌های مختلف فیزیکی و همچنین تمامی وضعیت‌های قرارگیری است. گذر از حالت Inter_Pose به Hard_Pose نیاز به یادگیری تغییر شکل دارد. تمامی دادگان توسط $n=3$ دوربین (زاویه دید) تولید می‌شوند و هر مجموعه داده شامل مجموعه‌ی آموزشی، اعتبارسنجی و تست است.

در این پژوهش از پایگاه داده‌ی Health [۲۸] برای ارزیابی اعتبارسنجی و تشخیص مفاصل بر اساس تشخیص ناهنجاری‌های اسکلتی استفاده شده است. اولین پایگاه داده جامع و در دسترس برای بررسی ناهنجاری‌های اسکلتی عضلانی با بیش از ۸۰۰۰ وضعیت قرارگیری سه بعدی برای رفتارهای مختلف مانند کار بر روی میز اداری در موقعیت ایستاده و نشسته به همراه تصاویر عمق و رنگی و دادگان اسکلتی که از دوربین کینکت جمع‌آوری شده است که جزییات آن در مرجع [۲۸] آمده است.

۳-۴ ارزیابی روش پیشنهادی

ارزیابی کارایی روش پیشنهادی در دو مرحله‌ی اصلی انجام می‌گیرد. در مرحله‌ی اول به بررسی تاثیر اعمال میزان اعتبار مفاصل بر نتایج کلاس‌بند می‌پردازیم و در مرحله‌ی دوم تخمین مفاصل از دست رفته مورد ارزیابی قرار می‌گیرد.

در این بخش جهت ارزیابی روش پیشنهادی از چهار معیار دقت^۱، حساسیت^۲، اختصاصیت^۳ و میانگین هندسی^۴ استفاده می‌شود. برای ارزیابی کلاس‌بند، میانگین معیارها از حلقه اجرایی (۱۰ مرتبه) 10-fold cross validation محاسبه می‌شود.

جدول ۲ نتایج بدست آمده از کلاس‌بندی داده‌های اسکلتی پایگاه داده Health به دو کلاس سالم و غیر سالم را با روش کلاس‌بند پیشنهادی WSVM بعد از اعتباربخشی و تخمین مفاصل از دست‌رفته در دو حالت نشسته و ایستاده نشان می‌دهد.

¹Accuracy

²Sensitivity

³Specificity

⁴G-mean

Pose اعمال می‌شود. در ابتدا شبکه با دقت ۷۸ درصد آموزش را شروع می‌کند و بعد از ۱۵۰,۰۰۰ تکرار با دقت ۸۲ درصد آموزش مرحله‌ی دوم پایانی می‌یابد. در ادامه نتایج بدست آمده از شبکه‌ی مرحله‌ی دوم بر روی مجموعه دادگان Hard-Pose اعمال می‌شود. در ابتدا شبکه با دقت ۷۳ شروع به آموزش می‌کند و بعد از ۸۸,۰۰۰ تکرار با دقت ۸۱ درصد آموزش پایان می‌یابد. نتایج اعلام شده در جدول ۴ نمایش داده شده است.

جدول ۴ مقایسه دقت کلاس‌بندی تراکم

Hard-Pose		Inter-Pose		Easy-Pose		مجموعه داده
پایان	شروع	پایان	شروع	پایان	شروع	
-	-	-	-	۸۷%	۰	شبکه ۱
-	-	۸۲%	۷۸%	۸۷%	۸۷%	شبکه ۲
۸۱%	۷۳%	۷۹%	۸۵%	۸۵%	۸۷%	شبکه ۳

بر اساس بخش‌بندی‌های انجام شده توسط شبکه CNN، 4 مفصل قابل پیش‌بینی نیستند. این ۴ مفصل شامل tip of the left hand، left thumb، tip of the right hand، right thumb هستند. نمودار شکل ۶ میزان کارایی الگوریتم تخمین مفاصل را بر اساس معیار دقت (Precision) در پایگاه داده جدول ۴ را نشان می‌دهد.

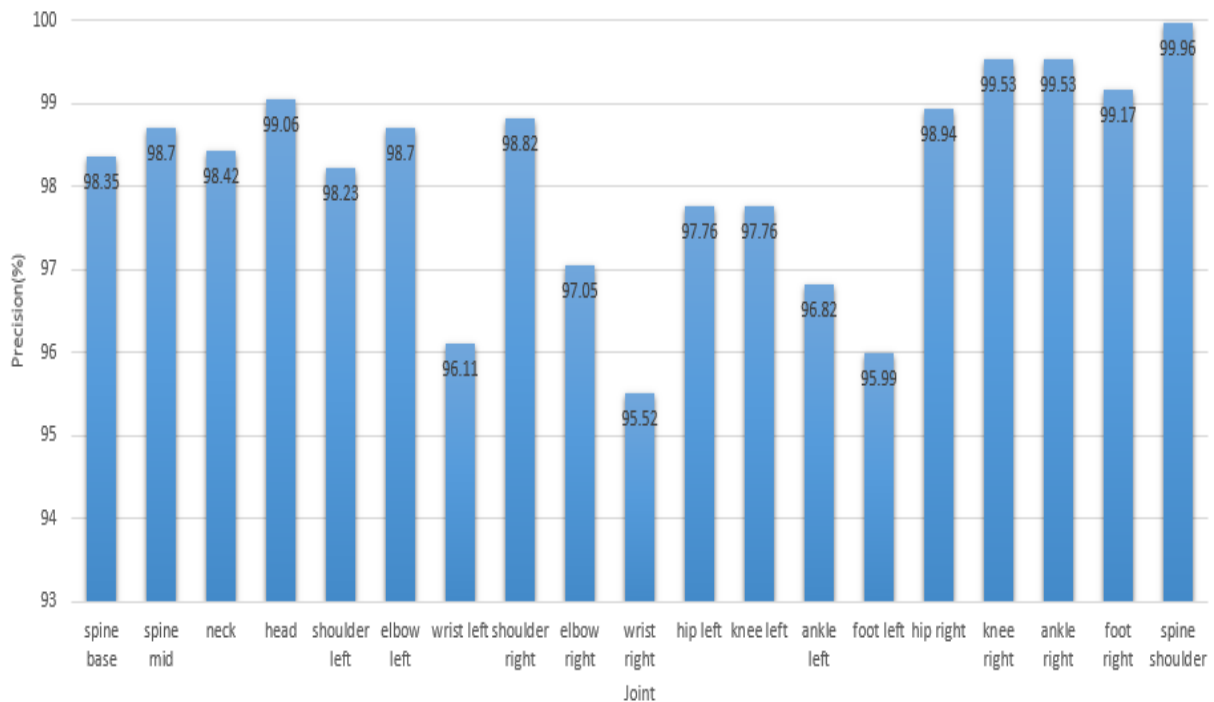
محلی از تصویر در نظر گرفته می‌شود که حساسیت فوق را برطرف می‌کند.

ویژگی‌های Color Correlogram نیز برخلاف خواص محلی مانند موقعیت پیکسل، جهت گرادیان و یا ویژگی‌های صرفاً سراسری مانند توزیع رنگ، همبستگی محلی و فضایی رنگی را به خوبی توزیع سراسری همبستگی فضایی در نظر می‌گیرد. در حالی که هر الگوریتم که بر اساس ویژگی‌های صرفاً محلی باشد به تغییرات بزرگ حساس است، Color Correlogram نسبت به این تغییرات پایدارتر است.

۴-۳-۱ ارزیابی تخمین مفاصل از دست‌رفته

در این بخش با اعمال حد آستانه‌ای برابر با مقدار ۰,۵، که بر اساس نتایج تجربی نتیجه‌ی بهتری را گزارش می‌دهد، ماتریس اعتبار بر اساس فریم عمق را مورد ارزیابی قرار داده و مفاصل با اعتبار پایین (مفاصل از دست‌رفته) را تشخیص داده و موقعیت آن‌ها تخمین زده می‌شود. سپس به بررسی تاثیر تخمین مفاصل از دست‌رفته بر روی کلاس‌بند پرداخته می‌شود.

همانطور که در بخش ۴-۱ گفته شد روند آموزش برای حالت‌های پیچیده Hard-pose طی سه مرحله انجام می‌گیرد. ابتدا شبکه با مجموعه دادگان Easy-Pose آموزش داده می‌شود. هر تکرار ۸ تصویر عمق را پردازش می‌کند و مرحله‌ی ابتدایی آموزش در ۲۵۰,۰۰۰ تکرار به دقت ۸۷/۸ درصد برای کلاس‌بندی تراکم بر روی مجموعه داده Easy-Pose دست می‌یابد. سپس نتایج بدست آمده از آموزش شبکه در مرحله اول بر حالت Inter-



شکل ۶ نمودار دقت مفاصل تخمین زده شده

- [6] Lara, O. D., Labrador, M. A., "A survey on human activity recognition using wearable sensors", IEEE Communications Surveys and Tutorials, vol. 15, pp. 1192-1209, 2013.
- [7] Machajdik, J., Hanbury, A., Garz, A., Sablatnig, R., "Affective computing for wearable diary and lifelogging systems: An overview", in Machine Vision-Research for High Quality Processes and Products-35th Workshop of the Austrian Association for Pattern Recognition. Austrian Computer Society, 2011.
- [8] Zordan V. B., Van Der Horst N. C., "Mapping optical motion capture data to skeletal motion using a physical model", in Proceedings of the 2003 ACM SIGGRAPH/Eurographics symposium on Computer animation, pp. 245-250, 2003.
- [9] Slyper, R., Hodgins, J. K., "Action capture with accelerometers", in Proceedings of the 2008 ACM SIGGRAPH/Eurographics Symposium on Computer Animation, pp. 193-199, 2008.
- [10] Shum, H. P., Komura, T., Takagi, S., "Fast accelerometer-based motion recognition with a dual buffer framework", The International Journal of Virtual Reality, vol. 10, pp. 17-24, 2011.
- [11] Kim, J. B., Kim, H. J., "Efficient region-based motion segmentation for a video monitoring system", Pattern Recognition Letters, vol. 24, pp. 113-128, 2003.
- [12] Felzenszwalb, P. F., Huttenlocher, D. P., "Pictorial structures for object recognition", International journal of computer vision, vol. 61, pp. 55-79, 2005.
- [13] Gupta, A., Kembhavi, A., Davis, L. S., "Observing human-object interactions: Using spatial and functional compatibility for recognition", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 31, pp. 1775-1789, 2009
- [14] Fouhey, D. F., Delaitre, V., Gupta, A., Efros, A. A., Laptev, I., Sivic, J., "People watching: Human actions as a cue for single view geometry", International journal of computer vision, vol. 110, pp. 259-274, 2014.
- [15] Chen, L., Wei, H., Ferryman, J., "A survey of human motion analysis using depth imagery", Pattern Recognition Letters, vol. 34, pp. 1995-2006, 2013.
- [16] Shotton, J., Sharp, T., Kipman, A., Fitzgibbon, A., Finocchio, M., "Real-time human pose recognition in parts from single depth images", Communications of the ACM, vol. 56, pp. 116-124, 2013.
- [17] Shotton, J., Girshick, R., Fitzgibbon, A., Sharp, T., Cook, M., Finocchio, M., et al., "Efficient human pose estimation from single depth images", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, pp. 2821-2840, 2013.
- [18] Ye, M., Wang, X., Yang, R., Ren, L., Pollefeys, M., "Accurate 3d pose estimation from a single depth image", in Computer Vision (ICCV), 2011 IEEE International Conference on, pp. 731-738, 2011.
- [19] Baak, A., Müller, M., Bharaj, G., Seidel, H. P., Theobalt, C., "A datadriven approach for real-time full body pose

۵ نتیجه‌گیری

در این پژوهش روشی برای تشخیص و تخمین موقعیت مفاصل دست رفته در دادگان اسکلتی حسگر کینکت ارائه شده است. همانطور که در بخش اول مطرح شد یکی از اهداف اصلی محیط‌های هوشمند و سیستم‌های نظارتی بهبود کیفیت زندگی انسان است. در واقع یک سیستم هوشمند با درک اینکه کاربر چه کاری را انجام می‌دهد و یا چگونه آن کار را انجام می‌دهد می‌تواند پشتیبانی مناسبی را ارائه دهد. مسائل بسیاری در زمینه‌های کاربردی گوناگون، دادگان اسکلتی را به عنوان ورودی سیستم به طور مستقیم استفاده می‌کند، همانطور که ذکر شد دادگان اسکلتی به دلیل چالش‌های موجود از اعتبار کافی برای سیستم‌های هوشمند برخوردار نیستند. از این رو نیاز به اعتباربخشی و اعتبارسنجی مطرح می‌شود.

در این مقاله نشان داده شد که با استفاده از توصیف‌گرهای مناسب برای تحلیل تصاویر می‌توان با دقت بالایی میزان اعتبار هریک از مفاصل را تعیین کرد و با شرکت دادن اعتبار در الگوریتم کلاس‌بندی دقت نتایج حاصل را بهبود بخشید، به طوری که با بهره بردن از میزان اعتبار، ویژگی‌های متمایز کننده‌ای برای الگوریتم کلاس‌بند ارائه داد. بنابراین ترکیب استخراج ویژگی متمایز کننده و الگوریتم کلاس‌بند پیشینه-حاشیه نتیجه‌ی خوبی را به همراه دارد. از طرف دیگر با استفاده از الگوریتم‌های یادگیری عمیقی همچون CNN می‌توان موقعیت مفاصل را با دقت خوبی تخمین زد و در فریم‌های چالشی به میزان قابل قبولی میانگین ماتریس اعتبار را بهبود بخشید.

مراجع

- [1] Weichert, F., Bachmann, D., Rudak, B., Fisseler, D., "Analysis of the accuracy and robustness of the leap motion controller", Sensors, vol. 13, pp. 6380-6393, 2013.
- [2] Shiratori, T., Park, H. S., Sheikh, Y., Hodgins, J. K., Motion capture from body mounted cameras, Google Patents, 2014.
- [3] Werghi, N., "Segmentation and modeling of full human body shape from 3-D scan data: A survey", IEEE Transactions on Systems, Man, and Cybernetics, Part C, vol. 37, pp. 1122-1136, 2007.
- [4] Boyali, A., Kavakli, M., Twamley, J., "Real time six degree of freedom pose estimation using infrared light sources and wiimote IR camera with 3D TV demonstration", in International Conference on Mobile and Ubiquitous Systems: Computing, Networking, and Services, pp.137-148, 2010.
- [5] Liu, Z., Zhu, J., Bu, J., Chen, C., "A survey of human pose estimation: the body parts parsing based methods", Journal of Visual Communication and Image Representation, vol. 32, pp. 10-19, 2015.

- [34] Bengio, Y., Louradour, M., Collobert, R., Weston, J., "Curriculum learning", presented at the Proceedings of the 26th Annual International Conference on Machine Learning, Montreal, Quebec, Canada, 2009.
- [35] Wang, J., Liu, Z., Wu, Y., Yuan, J., "Mining actionlet ensemble for action recognition with depth cameras", in Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on, pp. 1290-1297. 2012



عطیه‌سادات میرمعینی در سال ۱۳۹۷ کارشناسی ارشد خود را در رشته مهندسی کامپیوتر گرایش هوش مصنوعی از دانشگاه بوعلی سینا اخذ نمود. از زمینه‌های تحقیقاتی مورد علاقه ایشان می‌توان به هوش مصنوعی، بازشناسی الگو، پردازش تصویر، بینایی ماشین، و یادگیری عمیق اشاره نمود.



حسن ختن‌لو دانشیار گروه مهندسی کامپیوتر دانشگاه بوعلی سینا، در سال ۱۳۸۷ دکتری مهندسی کامپیوتر خود را از دانشگاه پیر و ماری کوری اخذ و تا به حال به عنوان عضو هیئت علمی گروه مهندسی کامپیوتر، مشغول به فعالیت می‌باشد و زمینه‌های تحقیقاتی مورد علاقه ایشان، پردازش تصویر و ویدیو، پردازش تصاویر و سیگنال‌های پزشکی، سیستم‌های فازی، شناسایی الگو و رباتیک و اتوماسیون صنعتی است.



وحید پورامین مربی گروه مهندسی کامپیوتر دانشگاه سید جمال‌الدین اسدآبادی، در سال ۱۳۹۰ مدرک کارشناسی ارشد خود را در رشته مهندسی کامپیوتر-گرایش هوش مصنوعی از دانشگاه بوعلی سینا اخذ نمود. وی از سال ۱۳۹۰ تا به حال به عنوان عضو هیئت علمی گروه مهندسی کامپیوتر، مشغول به فعالیت بوده و زمینه‌های تحقیقاتی مورد علاقه ایشان، هوش مصنوعی، بازشناسی الگو، پردازش تصویر، بینایی ماشین، یادگیری عمیق است.



الهام علیقارداش مربی گروه مهندسی کامپیوتر دانشگاه سید جمال‌الدین اسدآبادی، در سال ۱۳۹۱ مدرک کارشناسی ارشد خود را در رشته مهندسی کامپیوتر-گرایش هوش مصنوعی از دانشگاه بوعلی سینا اخذ نمود. وی از سال ۱۳۹۱ تا به حال به عنوان عضو هیئت علمی گروه مهندسی کامپیوتر، مشغول به فعالیت بوده و زمینه‌های تحقیقاتی مورد علاقه ایشان، هوش مصنوعی، بازشناسی الگو، پردازش تصویر، بینایی ماشین، یادگیری ماشین و یادگیری عمیق است.

- reconstruction from a depth camera", in Consumer Depth Cameras for Computer Vision, ed: Springer, pp. 71-98, 2013.
- [20] Kean, S., Hall, J., Perry, P., *Meet the Kinect: An introduction to programming natural user interfaces*, Apress, 2011.
- [21] Farooq, A., Jalal, A., Kamal, S., "Dense RGB-D Map-Based Human Tracking and Activity Recognition using Skin Joints Features and Self-Organizing Map", KSII Transactions on Internet & Information Systems, vol. 9, 2015.
- [22] Obdržálek, Š., Kurillo, G., Ofli, F., Bajcsy, R., Seto, E., Jimison, H., et al., "Accuracy and robustness of Kinect pose estimation in the context of coaching of elderly population", in Engineering in medicine and biology society (EMBC), 2012 annual international conference of the IEEE, pp. 1188-1193, 2012.
- [23] Alnowami, M., Alnowaimi, B., Tahavori, F., Copland, M., Wells, K., "A quantitative assessment of using the Kinect for Xbox360 for respiratory surface motion tracking", in SPIE Medical Imaging, pp. 83161T-83161T-10., 2012
- [24] Shum, Hubert H. P., Ho, S. L. Emdond, Jiang, Y., Takagi, S., "Real-time posture reconstruction for Microsoft Kinect", IEEE transactions on cybernetics, vol. 43, pp. 1357-1369, 2013.
- [25] Shafaei, A., Little, J. J., "Real-Time Human Motion Capture with Multiple Depth Cameras", in Computer and Robot Vision (CRV), 2016 13th Conference on, pp. 24-31, 2016.
- [26] Farnoosh, A., Ali, N. "Accurate body-part reconstruction from a single depth image", Multimedia Systems, vol. 25, pp. 165-176, 2019.
- [27] Liu, B., Li, Y., Zhang, S., Ye, X., "Healthy human sitting posture estimation in RGB-D scenes using object context", Multimedia Tools and Applications, vol. 76, pp. 10721-10739, 2017.
- [28] Ho, E. S., Chan, J. C., Chan, D. C., Shum, H. P., Cheung, Y. M., Yuen, P. C., "Improving posture classification accuracy for depth sensorbased human activity monitoring in smart environments", Computer Vision and Image Understanding, vol. 148, pp. 97-110, 2016.
- [29] European Agency for Safety and Health at Work, E-fact 45 - checklist for preventing bad working postures. <https://osha.europa.eu/en/publications/e-facts/efact45>.
- [30] Huang, J., Zabih, R., "Combining color and spatial information for content-based image retrieval", in Proceedings of ECDL, 1998.
- [31] Cantu-Paz, E., "Feature subset selection, class separability, and genetic algorithms", in Genetic and Evolutionary Computation-GECCO pp. 959-970. 2004
- [32] Charaoui, A. A., Padilla-López, J. R., Climent-Pérez, P., Flórez-Revuelta, F., "Evolutionary joint selection to improve human action recognition with RGB-D devices", Expert systems with applications, vol. 41, pp. 786-794, 2014.
- [33] Long, J., Shelhamer, E., Darrell, T., "Fully convolutional networks for semantic segmentation", in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3431-3440, 2015.