

ارائه روشی برای تحلیل و طراحی فضای بازی با پیش‌بینی نحوه نگاه بازیکن با بکارگیری شبکه‌های عصبی کانولوشنی

محمد درگاهی قره باغ^۱، یونس سخاوت^۲

^۱ کارشناسی ارشد هنرهای رایانه‌ای، دانشکده چندرسانه‌ای، دانشگاه هنر اسلامی تبریز، تبریز، ایران.
^۲ استادیار گروه چندرسانه‌ای، دانشکده چندرسانه‌ای، دانشگاه هنر اسلامی تبریز، تبریز، ایران.

نام نویسنده مسئول:

محمد درگاهی قره باغ

تاریخ دریافت: ۱۳۹۸/۱۱/۲۳

تاریخ پذیرش: ۱۳۹۹/۲/۶

چکیده

هدف از مطالعه نحوه طراحی بازی به دست آوردن خودکار آمدی و هدف‌مندی در مراحل طراحی آن است. یکی از جنبه‌های طراحی پیشرفت در خلق سطوح مؤثر در بازی است. مسئله‌ی اساسی در طراحی این است که بازیکن ضمن حفظ تعامل خود بتواند با ترکیبی از مهارت‌های آموخته شده تمرین کند. اگرچه مطالعاتی در رابطه با طراحی خودکار مراحل بازی انجام شده اما همچنان طراحان به دنبال روش‌های هدف‌مند برای تحلیل کار خود هستند. در این پژوهش روشی برای تحلیل و بررسی طراحی انجام شده بازی ارائه شده است. در روش ارائه شده مراحل بازی توسط مطالعه رفتارهای چشم انجام گرفته است. رفتارهای چشم در تحقیقات اخیر نتایج ارزشمندی در زمینه‌های روانشناسی و پزشکی ارائه کرده‌اند. این رفتارها توسط دستگاه ردیاب چشم استخراج می‌شوند. روال سخت و زمان بر این نتایج باعث بررسی روشی کارآمدتر برای استخراج رفتارهای چشم توسط شبکه‌های عصبی کانولوشنی شده است که در این پژوهش مورد استفاده قرار گرفته‌اند. با استفاده از این سیستم می‌توان بدون نیاز به دستگاه ردیاب چشم و مخاطب خاص آن به نتایج خوبی در این زمینه رسید. در ادامه پژوهش نتایج ۵ شبکه عصبی کانولوشنی با نتایج ردیاب چشم مورد ارزیابی قرار گرفتند. نتایج این مقایسه که در دو نوع بازی اول شخص و سوم شخص و دوبعدی و سه بعدی بود با استفاده از سیستم مشابهت سنجی پیکسل به پیکسل صورت گرفت. نتیجه ۴ شبکه عصبی و ۳ معیار مختلف با نتیجه ردیاب چشم مقایسه شد. نتایج حاصل نشان می‌دهد که درصد شباهت در بازی‌های سه‌بعدی سوم شخص بالاترین در صد در میان دیگر حالت‌ها است. در ادامه مطالعه رفتارهای چشم همچنین جهش چشم با استفاده از ردیاب چشم اندازه‌گیری شد که نتایج برای بازی NFS میزان ۰.۲۶، سطح معناداری به دست آمد. این نتیجه رابطه مستقیم جهش چشم بازیکن با بازی‌های اول شخص سه‌بعدی را نشان می‌دهد.

واژگان کلیدی: ردیاب چشم، شبکه‌های عصبی کانولوشن، رفتارهای چشم، طراحی محیط بازی.

مقدمه

طراحی فضای بازی شامل بافت^۱، صداگذاری، پویانمایی^۲ و فیزیک بازی^۳ می‌شود. فیزیک بازی است که توصیف می‌کند چگونه وضعیت بازیکن با گذشت زمان تکامل می‌یابد. فیزیک بازی مراحل درگیری در سطوح مختلف را شامل می‌شود که نتیجه آن نحوه عملکرد بازیکن در وضعیت بازی و تحقق صوتی و تصویری بازی است [۱]. فیزیک بازی شامل اشیاء مدل سازی شده، فضاها و تمام محتویات دیگر بازی است. تحلیل طراحی فیزیک بازی در پروسه ساخت بازی بسیار اهمیت دارد چنان‌که دنیل کوک^۴ طراح بازی ایجاد یک مدل قابل آزمایش از مکانیک بازی^۵ را مساوی با فرصت‌های جدیدی برای تعادل در بازی، طراحی اصلی بازی و کاربرد گسترده‌تر طراحی بازی در زمینه‌های دیگر می‌داند [۲].

همچنین یکی از مشکلاتی که در زمینه‌ی طراحی فیزیک بازی‌ها وجود دارد نبود ارتباط مؤثر بین موضوع و طراحی تصویری محیط بازی است [۳]. لذا بررسی طراحی تصویری مکانیک بازی منجر به کارآمدی بیشتر بازی می‌شود. الگوهای طراحی تصویری بازی‌ها راه‌حلی هستند که با گذشت زمان توسعه یافته و تکامل یافته‌اند. آن‌ها طراحی‌های اولیه برای شروع نیستند بلکه آن‌ها طراحی مجدد و بازیابی مجدد طراحی‌های دیگر توسعه دهندگان در تلاش‌های متعدد خود در طول زمان هستند. طراحان این الگوهای طراحی را نمونه‌ای برای شروع طراحی فضا در نظر می‌گیرند [۴]. طراحان بازی برای تحلیل این فضاها نیاز به دانش بصری و سابقه طراحی فضاهای موفق از بازی هستند. هر چند موفقیت پی در پی در این زمینه دشوار به نظر می‌رسد. مطالعه [۵] تلاشی برای ایجاد یک چارچوب در طراحی بازی‌های آموزشی است. هدف این مطالعه آگاهی از نحوه مؤثر بودن مهارت‌ها و دانشی است که یک بازیکن نیاز است در طول بازی به آن‌ها نیاز داشته باشد و یا او را تشویق به جست و جوی بیشتر کند. همین‌طور از دیگر راه‌حل‌ها برای بررسی طراحی بازی‌ها بررسی رفتار چشم بازیکن در حین بازی است. پژوهش [۶] از این موضوع برای بررسی میزان رضایت، تعامل یا غوطه‌وری بازیکن استفاده کرده است. روش‌های ارائه شده برای آزمون بازی‌ها توسط این روش اجازه می‌دهد تا جنبه‌های دیگر استفاده بازی مورد بررسی قرار بگیرد و همچنین مراحل بازی و کشف دانش جدید در بازی بهبود یابد.

انجام فرایند تحلیل طراحی بازی‌ها مستلزم به کارگیری چارچوب نظری و عملی است. تحلیل مکانیزم اثرگذاری بازی بر افزایش مهارت بازیکن و تکامل مستمر بازی با استفاده از نتایج رفتار چشم باعث بروز چارچوب نظری می‌شود. استفاده از این چارچوب و طراحی‌های مؤثر پی در پی و جایگزین الگوی پیشین، یک الگوی موفق برای موضوع هدف را مشخص می‌کند. رفتارهای چشم همانند دیگر رفتارهای اندام انسان دلیلی بر واکنش آن نسبت به موضوع رو به رو است. خیرگی^۶ و جهش^۷ از جمله رفتارهای مشخص شده برای چشم هستند که هر کدام در شرایطی حاصل می‌شوند. برای مثال در زمانی که در تصویر اختلاف رنگ یا تضاد رنگی وجود داشته باشد و یا اگر تصویر دارای بعد باشد چشم به آن نقطه خیره خواهد شد [۷، ۸]. نقاط خیرگی دارای اهمیت بیشتری نسبت به دیگر نقاط تصویر در طراحی مرحله‌ای از بازی هستند. با در نظر گرفتن شرایط رفتاری چشم می‌توان قاعده‌ای هدف‌مند برای طراحی نقاط مورد هدف تمرکز و غیرتمرکز در نظر گرفت. با توجه به نکات ذکر شده نحوه استفاده از اشباع^۸ و نور^۹ رنگ‌ها در طراحی اشیاء و فضای بازی اهمیت پیدا خواهند کرد. تغییرات پی در پی در روند بازی و سیال بودن بازی‌ها امکان بروز احتمالات پیچیده‌تری در فضای بازی را می‌دهد که استفاده مداوم از دستگاه ردیاب چشم^{۱۰} برای تحلیل طراحی بازی و مخاطبان در دسترس را سخت می‌کند. از این‌رو هوشمند سازی این مکانیزم باعث افزایش سرعت در تحلیل، کاهش هزینه و زمان تحلیل بازی می‌شود.

¹ Texture

² Animation

³ Physical Layout

⁴ Daniel Cook

⁵ Model of Game Mechanics

⁶ Gaze

⁷ Saccade

⁸ Saturation

⁹ Lightness

¹⁰ Eye Tracking

هوشمندسازی استخراج رفتارهای چشم یا پیش‌بینی رفتارهای چشم توسط شبکه‌های عصبی کانولوشنی^{۱۱} انجام گرفته است. این نوع سیستم‌ها دارای عملکرد یادگیری از داده‌های واقعی استخراج شده در زمینه‌های مختلف هستند. داده‌های استخراج شده از دستگاه ردیاب چشم نیز به عنوان داده‌های ورودی آموزشی این سیستم‌ها استفاده شده است. در پژوهش‌های اخیر [۹] نتایج پیش‌بینی در مرحله آزمون این سیستم‌ها نتایج بسیار دقیق و نزدیک به نتیجه واقعی را نشان داده است.

پژوهش حاضر نتایج به دست آمده از دستگاه ردیاب چشم و شبکه‌های عصبی کانولوشنی را برای بررسی میزان بهره‌وری شبکه‌های عصبی مقایسه کرده است. نتایج دستگاه ردیاب چشم در نوع بازی اول شخص^{۱۲} و سوم شخص^{۱۳} و دو حالت دو بعدی و سه بعدی استخراج شد و همین‌طور نتایج رفتار چشم تصاویر بازی‌ها توسط چندین شبکه عصبی کانولوشنی پیش‌بینی شده است. در ادامه پژوهش میزان بهره‌وری نتایج روش ذکر شده بررسی شده است.

جهش چشم نوع دیگری از رفتار چشم است که در این پژوهش با هدف بهبود عملکرد طراحان بازی مورد ارزیابی قرار گرفته است. رفتارهای چشم طبق مطالعات پژوهش‌گران برای تعیین سلامتی و میزان هوشیاری آن‌ها در زندگی روزمره نیز بررسی شده است [۱۰، ۱۱]. در مطالعه [۱۲] برای شناسایی بیماران اسکیزوفرنیک از جهش چشم برای تولید بازی جدی برای این نوع بیماران استفاده شده است. در روند این تحقیقات مشاهده می‌شود که تحلیل رفتارهای چشم همانند جهش در تحلیل محیط بازی برای بررسی سلامتی بازیکن نیز می‌تواند حائز اهمیت بالایی باشد. در این نوع رفتار که رابطه‌ی مستقیمی با تمرکز و عدم تمرکز دارد می‌توان از آن برای هدف بهبود تمرکز بازیکن در طراحی بازی استفاده نمود.

روش پیشنهادی با چندین شبکه عصبی مورد آزمون قرار گرفت. نتایج مقایسه نتایج نشان داد که شبکه‌های عصبی کانولوشنی قابلیت تحلیل فضای بازی را دارند. همچنین با به وجود آوردن تغییراتی در ساختار معماری شبکه می‌توان عناصر مربوط به بازی را برای بررسی بهتر در پیش‌بینی نتایج به کار برد. تشخیص بعد، اول شخص یا سوم شخص بودن، بافت و حتی موسیقی بازی [۱۳] می‌تواند برای ساخت یک شبکه عصبی جدید برای تحلیل ویژه مکانیک بازی مورد پژوهش قرار گیرد.

در ادامه بخش‌های پژوهش به شکل زیر ساماندهی شده است: بخش ۲ به پژوهش‌های پیشین در رابطه با استفاده از شبکه‌های عصبی، ردیاب چشم و رفتارهای چشم پرداخته است. روش پیشنهادی و انواع شبکه‌های عصبی برای استفاده در طراحی بازی و تحلیل رفتارهای چشم در بخش ۳ توصیف شده‌اند و در بخش ۴ به ارزیابی نتایج و مقایسه نتایج پرداخته شده است. در نهایت بخش ۵ به نتیجه‌گیری و کارهای آتی پرداخته است.

پیشینه پژوهش

تحلیل رفتارهای چشم به دو روش انجام می‌شود. در روش اول با استفاده از دستگاه ردیاب چشم جهش چشم، خیرگی و نگاه مستمر^{۱۴} قابل اندازه‌گیری است. دستگاه‌های ردیاب چشم برای بررسی حرکت چشم فرد در محیط مجازی استفاده می‌شوند و اجازه حرکت آزادانه در فضای محیط واقعی را می‌دهند که یک مؤلفه مهم در تصمیم‌گیری و انتخاب محیط مناسب برای بازی رایانه‌ای است. از آنجایی که رفتار دیداری به‌وسیله‌ی نگاشت نقطه خیرگی بر روی تصویر مرجع انجام می‌شود با استفاده از یک سیستم ردیاب چشم جهت دید را در مختصات جهان سه‌بعدی ارزیابی می‌کنند [۷]. استفاده از ردیاب چشم به دو حالت دو بعدی و سه بعدی امکان پذیر است. نتایج حاصل از روند استخراج رفتارهای چشم به دو شکل نقشه گرمایی^{۱۵} و نقشه برجستگی^{۱۶} است. از کاربردهای این نوع تجزیه و تحلیل طراحی زیبایی شناسانه است. با استفاده از نقشه‌ی گرمایی، می‌توان گفت که کدام قسمت‌ها بیشتر به نقاط مورد توجه کاربر تعلق دارند. تجزیه و تحلیل زیبایی‌شناختی فعلی به تصاویر دوبعدی متکی است [۱۴]. از این‌رو بررسی زیبایی‌شناختی سه بعدی کمی دشوار به نظر می‌رسد.

¹¹ CNN

¹² FPS

¹³ TPS

¹⁴ Smooth Pursuit

¹⁵ Heat map

¹⁶ Saliency

خیرگی‌های چشم به‌دست‌آمده توسط ردیاب چشم به‌طور کلی به سه کلاس مختلف، مانند نگاه ثابت، جهش چشم و نگاه مستمر و پیوسته طبقه‌بندی می‌شوند [15]. در [۱۶] روشی برای مجزا کردن نگاه پیوسته از سایر رفتارهای چشم نیز [15] ارائه شده است. این نکته می‌تواند در بهبود تحلیل نتایج استفاده شود. در پژوهش [۱۷] اثربخشی رفتارهای جنبش چشم مرتبط با عناصر تبلیغاتی مورد مطالعه قرار گرفته است. در این پژوهش اثربخشی نحوه رفتار حرکات چشم به عناصر آگهی تبلیغات ویدئویی بررسی شده که نشان می‌دهد اثرات تبلیغاتی مختلف، ارتباط خاصی با رفتار حرکت چشم در آگهی دارند و همچنین تمرکز خود را بر مجموعه‌ای از نقاط خیرگی محدود کرده‌اند و با استفاده از یادگیری عمیق این خیرگی را در طول مشاهده دنبال کرده‌اند. استفاده گسترده از تحلیل رفتارهای چشم در زمینه‌های مختلف از کاربرد ثمربخش این رویکرد جدید است. در تحقیقات انجام گرفته نتیجه بر این شده است که خیرگی و نگاه ثابت^{۱۷} کاربر ۴۰ تا ۷۰ درصد به مرکز تصویر معطوف می‌شود که این مسئله اهمیت مرکز تصویر را نشان می‌دهد. روانشناسان از این موضوع به نام تعصب دید مرکزی^{۱۸} انسان نام می‌برند [۱۸]. این نوع تعصب بیان می‌کند که انسان همواره برای نخستین بار برای تصویر جدید ابتدا به مرکز تصویر نگاه می‌کند. این موضوع در یکی از پژوهش‌های شبکه‌های عصبی مورد بررسی قرار گرفت تا نتایج نهایی در تعادل با این موضوع باشند. در پژوهش [۱۹] تشخیص نگاه ثابت^{۱۹} با استفاده از مفهوم الگوریتم سطل چکه کن^{۲۰} [۲۰] انجام می‌شود، جایی که نقاط خیرگی دید بر روی هم افزونگی و اتصال پیدا کرده تا تشکیل خیرگی دهند. این نکته به ویژگی‌هایی از قبیل حرکت ثابت چشم، قطر مردمک، مدت زمان نگاه ثابت^{۲۱} اشاره دارد که از آن‌ها برای توسعه مدل‌های شناختی استفاده می‌شود.

جهش چشم به معنی حرکت سریع چشم است و حدود ۳۰ تا ۱۲۰ میلی‌ثانیه زمان می‌برد که بین تثبیت‌های چشم اتفاق می‌افتد. تثبیت‌ها نقاط تمرکز در تصویر هستند که حدوداً در ۰٫۲ تا ۰٫۸ ثانیه اتفاق می‌افتند [۲۱]. با این نوع رفتار چشم، ما تصاویری را که از جهش و تثبیت بیشتری برخوردارند تصاویری با اطلاعات و روشنایی‌های متفاوتی می‌بینیم. اطلاعات با نگاه کردن چشم، تمام شیء‌ها را در نقطه‌ی فوویا^{۲۲} تنظیم می‌کند تا نقاط تثبیت به وجود آید [۲۱]. در یک بازی ممکن است تعداد شیء‌ها زیاد باشد و چشم اکثراً در حال تغییر نقاط تثبیت شود و در نتیجه تعداد جهش‌های چشم بسیار افزایش یابد. در نتیجه چشم بازیکن بعد از مدتی بازی کردن خسته می‌شود. در مطالعات انجام شده بر روی این مباحث [۲۲] نتیجه‌گیری شده است که میزان پلک زدن مخاطب با تثبیت چشم رابطه عکس و میزان جهش چشم با قدرت بینایی چشم بازیکن رابطه مثبتی دارد. همچنین این پژوهش [۲۳] به رابطه بین رفتار بصری و انواع مختلف حالت‌های روحی مانند کنجکاوی [۲۴]، مهارت [۲۵]، و کارایی [26] هم اشاره کرده است.

در پژوهش [۲۷] از ویژگی‌های چشم در بازی‌های جدی برای اندازه‌گیری میزان توجه کودکان بهره گرفته‌اند. توسط این پژوهش می‌توان نتایج به دست آمده از خیرگی‌ها را برای یادگیری شبکه جدید استفاده کرد تا نتیجه‌ی افزایش توجه را در بازی‌های دیگر مورد بررسی قرار دهد. همچنین برای حالات بازیکن در پیش‌نمایش بازی بررسی شده است. در ادامه این پژوهش بازیکن انتظار دارد در هنگام بازی همچون تجربه پیش‌نمایش تعامل با بازی برای او جذابیت اولیه را داشته باشد [۲۸].

ردیاب چشم برای بررسی رفتارهای چشم بازیکنان در بازی Tomb Rider برای پنج شرکت کننده مورد آزمون قرار گرفت [۶]. در طول این آزمون مدت زمان تثبیت چشم برای بازیکن اندازه‌گیری شده است. هر چند تعداد شرکت کننده برای انجام آزمون اندک است. در این رویکرد حالت بازی و احساس بازیکن نسبت به مکانیزم بازی با استفاده از کلمات کمی مانند ساده، خسته کننده و غیره سنجیده شده است. نتایج حاصل از این پژوهش نشان می‌دهد ردیاب چشم مشکلاتی همچون عدم تعامل بازیکن با اشیاء تعاملی را تشخیص می‌دهد. در پژوهش حاضر علاوه بر نکات پژوهش پیش به ابعاد بازی‌ها، جهش چشم و اول شخص و سوم شخص بودن بازی‌ها با تعداد بیشتری از آزمون شوندگان پرداخته شده است.

¹⁷ Fixation

¹⁸ Central Fixation Bias

¹⁹ Fixation Detection

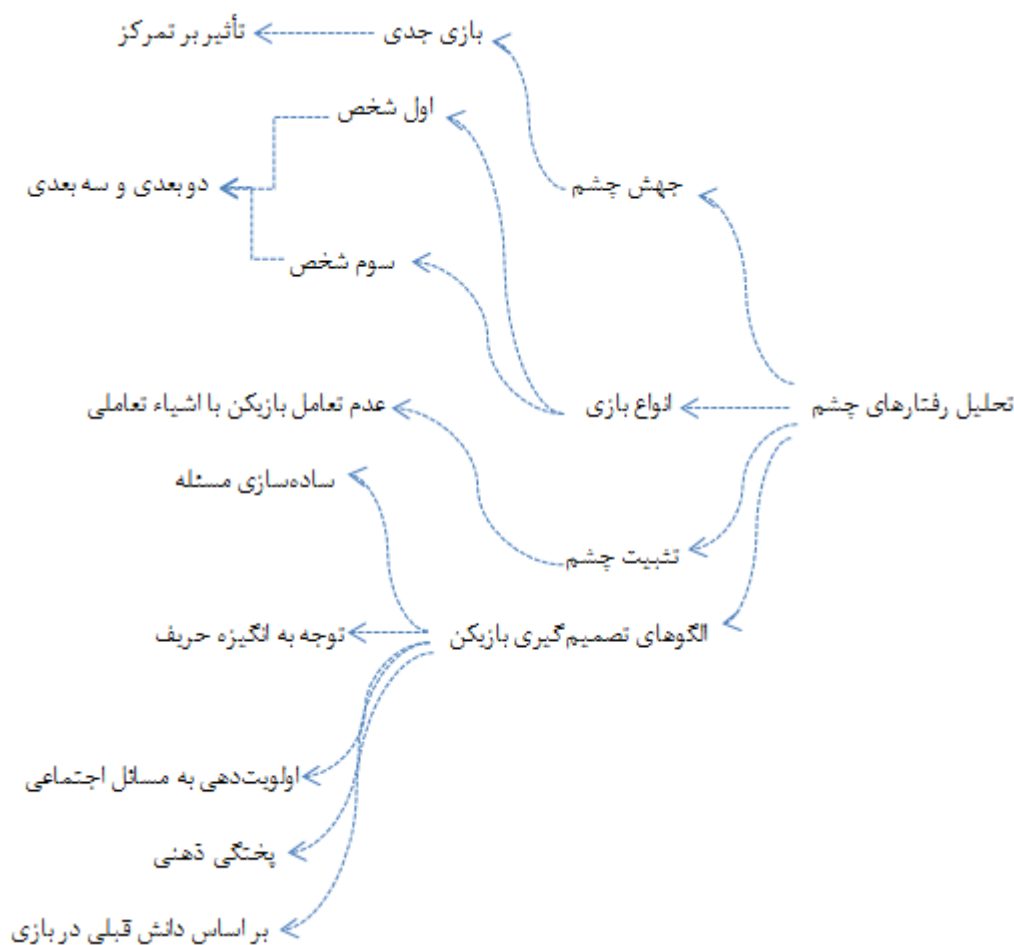
²⁰ Leaky Bucket

²¹ Fixation Durations

²² Fovea

در پژوهش [۲۹] به بررسی نحوه تصمیم‌گیری بازیکن در انواع ژانر بازی با استفاده از ردیاب چشم پرداخته شده است. نتایج این پژوهش الگوهای جداگانه اما پایدار را برای طراحی تصویری بازی‌ها نشان می‌دهد که مبتنی بر پختگی ذهنی بازیکن و اولویت‌دهی به مسائل اجتماعی بود. موضوع ژانر و نحوه تصمیم‌گیری در بازی در پژوهش حاضر مورد بررسی قرار نگرفته است. همچنین در [۳۰] بازیکنان از تصمیم‌گیری کاملاً منطقی استفاده می‌کنند که شامل بهترین پاسخگویی به ساده‌سازی مسئله تصمیم، یا با نادیده گرفتن انگیزه‌های بازیکن مقابل یا در نظر گرفتن آن‌ها فقط برای زیرمجموعه نتایج است و در آخر با مشاهده انواع حرکات چشم و انتخاب بازیکن در بازی ارتباطی بین آن‌ها پیدا می‌کنند.

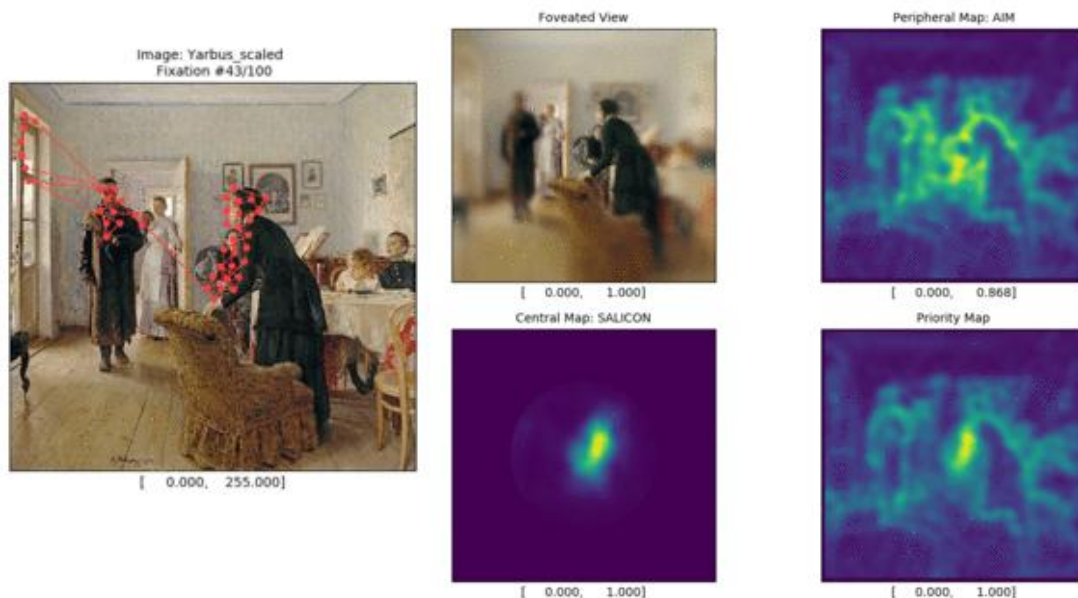
تصویر ۱ خلاصه‌ای از تأثیر رفتارهای چشم بر جنبه‌های مختلف بازی را نشان می‌دهد.



تصویر ۱. خلاصه‌ای از تأثیرات رفتارهای چشم بر انواع جنبه‌های بازی

روش دوم استخراج رفتارهای چشم با استفاده از شبکه‌های عصبی کانولوشنی است. این نوع سیستم با استفاده از یادگیری داده‌های استخراج شده توسط ردیاب چشم با تعداد آزمون‌شوندگان زیاد انجام می‌شود. در این حالت داده‌های برجستگی [۳۱]، که نوعی خروجی با فرمت تصویر دیجیتال است به عنوان داده ورودی شبکه و برای آموزش مدل شبکه استفاده می‌شود. از این‌رو حالت داده‌های بسیار مؤثر است. داده‌هایی که برای تصاویر بازی‌های ویدیویی انجام گرفته با داده‌های تصاویر پزشکی متفاوت است و نتایج پیش‌بینی متفاوتی را خواهند داشت. بنابراین برخی از سیستم‌های پیش‌بینی رفتارهای چشم با چند مجموعه داده آموزش می‌بینند. در پژوهش [۳۳] از چندین مجموعه داده مانند GBVS و Itti & Koch استفاده شده است. در مجموعه داده‌ای که این پژوهش طراحی کرده است از ۱۵ آزمون‌شونده برای ۷۰۰ تصویر با عنوان‌های مختلف نتایج ردیاب چشم را استخراج کرده‌اند. همچنین در این پژوهش سه مرحله تشخیص اشیاء، بخش‌بندی اشیاء و ویژگی‌های سطح پایین تصاویر توسط لایه‌های کانولوشنی استخراج می‌شود. این مراحل باعث افزایش دقت نتایج پیش‌بینی در انواع تصاویر با ویژگی

مختلف می‌شود. همچنین این موضوع ضرورت استفاده از داده‌های مخصوص بازی را برای نتایج بهتر در زمینه بازی گوش‌زد می‌کند. بازه سنی زیاد شرکت کنندگان در پژوهش پیش باعث بهبود نتایج در انواع تصاویر در ژانرهای متفاوت باشد اما تعداد اندک شرکت کنندگان این نکته را دچار ابهام می‌کند. پژوهش [۳۴] با انجام تحقیقاتی نشان می‌دهد که اکثر مدل‌های پیش‌بینی برای انواع تصاویر نتایج خوبی ندارند و چگونه می‌توان شبکه درستی برای نتایج درست در تمام زمینه‌ها ساخت. اکثر مدل‌های پیش‌بینی رفتارهای چشم با نتایج MIT300 و LSUN بررسی می‌شوند در پژوهش پیش با متداول‌ترین معیارهای اندازه‌گیری (AUC, sAUC, NSS, CC, SIM, KL-Div) نتایج اندازه‌گیری می‌شوند. نتایج نشان می‌دهد که نمی‌توان هر مدلی را برای زمینه‌های مختلف به کار بست اما یک مدل خوب در بیشتر زمینه‌ها به درستی عمل می‌کند. در پژوهش [۹] علاوه بر یادگیری مدل از داده‌های ورودی برای تعصب دید مرکزی سیستم یادگیری چندباره^{۲۳} طراحی شده است تا تأثیر آن بر نتایج خروجی خنثی شود. اکثر شبکه‌های عصبی برای پیش‌بینی خیرگی و تثبیت چشم طراحی شده‌اند و اندک پژوهشی در رابطه با جهش چشم ساخته شده است. در پژوهش اخیر [۳۵] شبکه‌ای برای پیش‌بینی جهش چشم ارائه شده است. این تحقیق یک مدل تولیدکننده چند جهشی مبتنی بر ادغام سطح برجستگی مرکزی و مبتنی بر شیء و همچنین ویژگی‌های سطح پایین برجستگی ایجاد می‌کند که با بانک اطلاعاتی CAT2000 ارزیابی می‌شود. نمونه‌ای از نتایج جهش چشم را برای نقاشی یاربوس در تصویر ۲ قابل مشاهده است همچنین تصویر برجستگی برای همین نقاشی در سمت چپ قرار دارد.



تصویر ۲. سمت راست: جهش چشم با نقاط قرمز برای نقاشی یاربوس. سمت چپ: نقشه برجستگی برای همان نقاشی

در پژوهش اخیر [۳۶] برخلاف سیستم‌های قبلی بر طیف گسترده‌ای از خانواده مدل‌های CNN قابل استفاده است: CNN با لایه‌های کاملاً متصل (مانند VGG)، CNN‌هایی برای خروجی‌های ساختارمند (مانند زدن زیرنویس^{۲۴})، CNN‌هایی با ورودی‌های چند حالتی (برای مثال جواب به سؤالات بصری^{۲۵}) یا در یادگیری‌های تقویتی می‌توانند بدون یادگیری دوباره و تغییر در معماری شبکه استفاده شود. این مدل با ترکیب داده‌های موجود با مدل ساختاری خود به کلاس‌های جداکننده بصری با وضوح بسیار بالا از نظر نتایج دست پیدا می‌کند که همگی با طبقه‌بندی^{۲۶} تصاویر، زیرنویس کردن تصاویر و پاسخ‌دهی به سؤالات بصری پیوند داده می‌شوند. در واقع این سیستم با جست و جوی محلی اشیاء در تصویر ورودی دلیلی برای

²³ Learned Priors

²⁴ Captioning

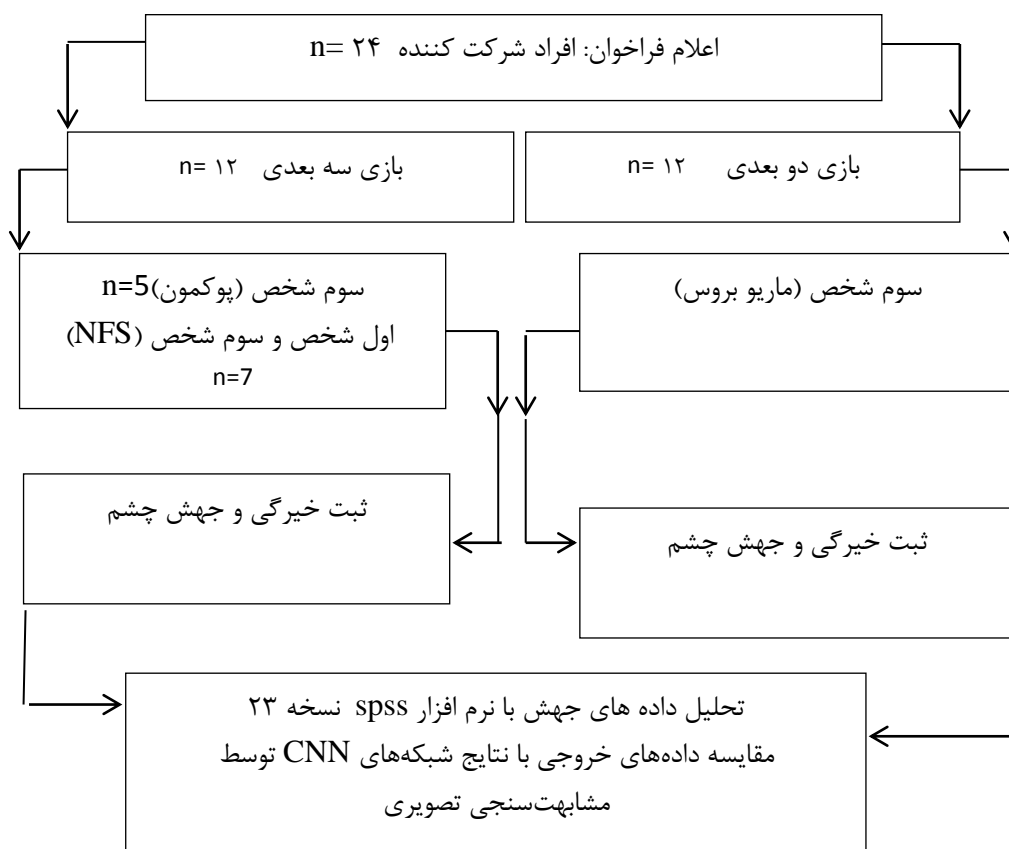
²⁵ Visual Question Answering (VQA)

²⁶ Classification

پاسخ به چرایی به وجود آمدن تصویر پیدا می‌کند. این نتیجه بسیار قابل استفاده در روایی‌سازی داستان بازی‌ها است به این شکل که مکان‌های اشیاء استفاده شده در طراحی همگی تفسیری برای به کار برده شدن در بازی خواهند داشت.

روش پیشنهادی

برای انجام آزمون و استخراج داده‌های ردیاب چشم ابتدا اطلاعاتی درج شد که داوطلبان به‌منظور انجام این آزمون مراجعه کنند. در نهایت آزمون‌شوندگان از دانشجویان دختر و پسر دانشگاه هنر اسلامی تبریز به صورت تصادفی انتخاب شدند. در ادامه ۲۴ نفر به‌منظور این آزمون انتخاب شدند که استعداد انجام بازی‌های رایانه‌ای را داشتند. قبل از آزمون توضیحاتی در رابطه با نحوه عملکرد و روند گرفتن آزمون داده شد تا اضطراب‌های بی‌مورد برای آزمون‌شونده از بین برود. آزمون با بازی ماریو بروس برای بازی دو بعدی انجام شد. تعداد ۱۲ نفر (۶ نفر مؤنث و ۶ نفر مذکر) این بازی را انجام دادند و داده‌های آن‌ها ثبت گردید. سپس استخراج داده‌های بازی پوکمون که یک بازی سه‌بعدی و اول‌شخص بود با ۵ آزمودنی (۳ نفر مذکر و ۲ نفر مؤنث) انجام گرفت. در این آزمون علاوه بر ثبت داده‌های خیرگی، جهش چشم نیز ثبت شد. دستگاهی که برای آزمون استفاده شد دستگاه Tobii 4c است. همچنین برای ذخیره داده‌ها از نرم‌افزار توبی گیزویور استفاده شد که برای این پروژه از جهت ثبت جهش چشم مناسب بود. این نرم‌افزار قابلیت ایجاد نقشه‌ی حرارتی دارد که برای مقایسه با تصویر خروجی از شبکه عصبی اهمیت داشت. خروجی از نرم‌افزار توبی گیزویور به ۳ حالت ذخیره شد که یک حالت برای نوع خام فیلم بازیکن برای استفاده ورودی شبکه‌های عصبی است. داده‌های به‌دست‌آمده در آدرس آزمایشگاه واقعیت افزوده شناختی دانشگاه هنر اسلامی تبریز (www.carlab.ir) در دسترس هستند. نمودار ۱ روند گرفتن آزمون را مشخص کرده است.



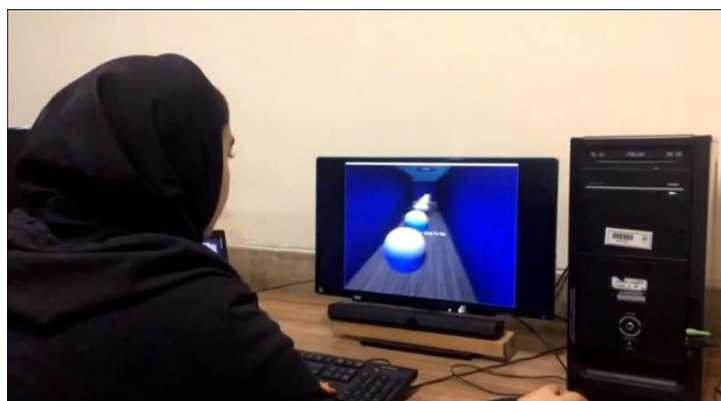
نمودار ۱. روند آزمون

تعدادی از نتایج ۳ بازی استفاده شده توسط آزمون شوندگان در تصویر ۳ آمده است.



تصویر ۳ دو عکس سمت راست بازی NFS و دو تصویر سمت چپ ماریوبروس و پوکمون

تصویر ۴ نیز آزمون شونده را در حال بازی نشان می‌دهد.



تصویر ۴ آزمون شونده در حال بازی پوکمون

در ادامه پژوهش ۵ شبکه عصبی کانولوشنی که توضیحات آن‌ها در بخش ۲ آمد مورد استفاده قرار گرفتند. خلاصه‌ای از معماری و شبکه این ۵ شبکه عصبی در جدول ۱ مشخص شده است. نمودار ۲ خلاصه‌ای از کاربرد های ۵ شبکه عصبی استفاده شده در پژوهش را مشخص کرده است.

- GBVS ← ۹۸ درصد دقت بر روی ROC^{۲۷} ← ۷۴۹ تصویر با ۱۰۸ طبقه‌بندی مختلف [۳۷]
- Itti & Koch ← ۸۴ درصد دقت بر روی ROC [38-40]
- AUC, NSS, CC ← اندازه گیری نتایج با معیارهای مختلف ← ۱۰۰۰ مجموعه ۱۰۰ عکسی با نقشه تثبیت [۳۴]
- Grad-CAM ← تفسیر و پاسخ به نتایج بصری ← زیرنویس عکس، لایه‌های کاملا متصل، خروجی ساختارمند [۴۱]
- DCNN ← شبکه‌های عصبی کانولوشنی عمیق ← برقراری ارتباط بین روش مصورسازی DCNN^{۲۸} و deCON^{۲۹} [۴۲]

نمودار ۲ خلاصه‌ای از کاربردهای ۵ شبکه عصبی استفاده شده در پژوهش

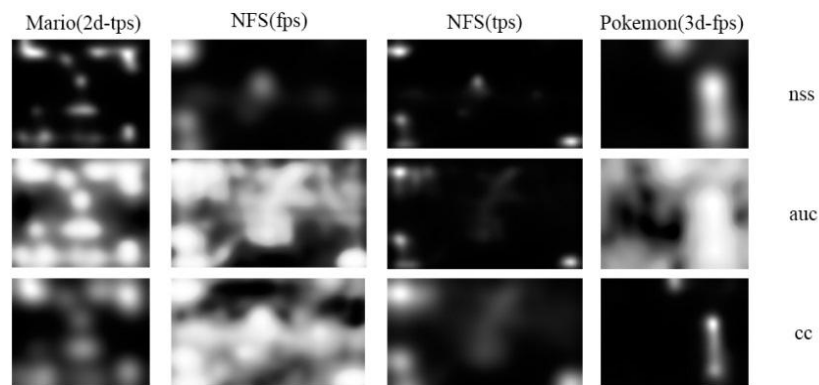
ابعاد بازی رابطه مستقیمی با اول شخص بودن و سوم شخص بودن بازی دارد. بازی‌های دو بعدی امکان ایجاد بعد در حالت سه بعدی را ندارند. در این پژوهش سه بازی مختلف NFS، ماریوبروس و پوکمون انتخاب شده‌اند. بازی NFS یک بازی هیجانی و پرشور است که کاملا نیاز به تمرکز دارد. بازی ماریو یک بازی قدیمی و پرترفدار است که همچنان نسخه‌های جدیدتر

²⁷ Receiver operating characteristic

²⁸ Deep Convolutional neural Network

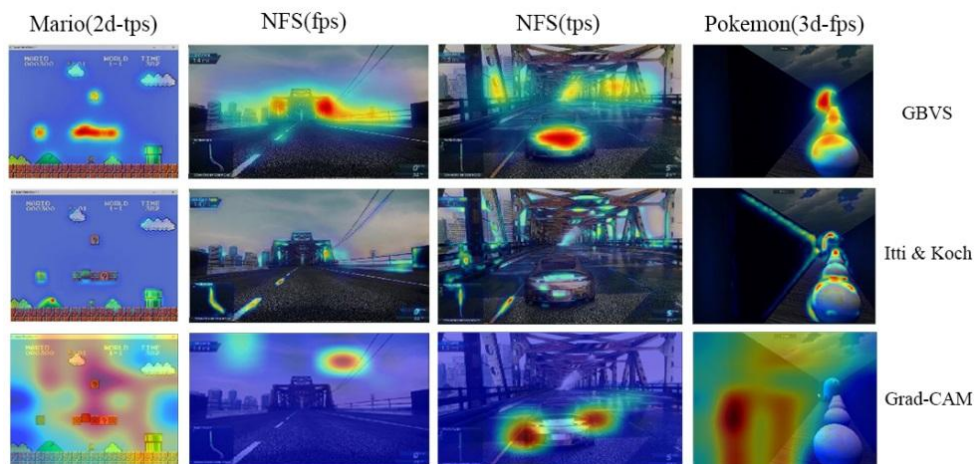
²⁹ Deconvolution

از آن ساخته می‌شود. بازی پوکمون در حالت سه‌بعدی و اول شخص در اختیار آزمون‌شوندگان قرار گرفت. همچنین نتایج برجستگی این سه بازی در تصویر ۵ برای سه معیار AUC، NSS و CC مشخص شده است.



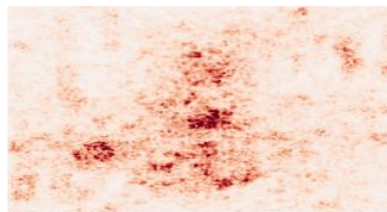
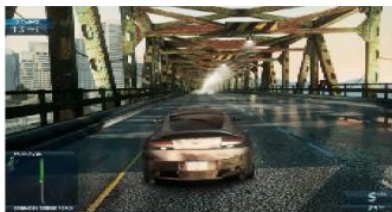
تصویر ۵ نتایج برجستگی برای ۳ بازی در ۳ معیار AUC، NSS و CC

نتایج بالا در مقایسه دو معیار NSS و CC پیش‌بینی نسبتاً نزدیکی دارند در حالی که نتایج معیار AUC نسبت به دو نتیجه دیگر قسمت‌های دیگری از تصویر را انتخاب کرده است. برخی از نتایج بصورت برجستگی هستند که در تصویر ۵ مشاهده می‌شود و برخی از نتایج برای شبکه‌های GBVS، Itti & Koch و Grad-CAM به شکل نقشه حرارتی هستند. در تصویر ۶ نتایج سه شبکه مورد بررسی قرار گرفته است.



تصویر ۶. نتایج سه شبکه عصبی GBVS، Itti & Koch و Grad-CAM

همان‌طور که در تصویر ۶ قابل مشاهده است نتایج دو شبکه GBVS و Itti & Koch تفاوت نسبتاً کمی در محل تمرکز و خیرگی نسبت به یکدیگر دارند در حالی که تفاوت این دو نسبت به نتیجه Grad-CAM محسوس است. دو شبکه اول برای بررسی مهمترین شیء تصویر تلاش می‌کنند و خیرگی نگاه مخاطب را در تصویر ورودی پیش‌بینی می‌کنند در حالی که نتایج شبکه سوم برای بررسی تفسیر و پاسخ به چرایی تصویر ورودی استفاده می‌شود. در واقع پیش‌بینی شبکه عصبی Grad-CAM دلیلی برای بوجود آمدن تصویر را نشان می‌دهد. یکی دیگر شبکه‌های مقایسه در این مقاله با نام DCNN نام گذاری شده که به دلیل برقراری ارتباط بین شبکه‌های عمیق و Deconvolution است. نتیجه برجستگی متفاوتی از این شبکه از لحاظ ظاهری نسبت به نتایج پیش در تصویر ۷ مشاهده می‌شود. این شبکه نیز برای پیش‌بینی خیرگی استفاده می‌شود.



تصویر ۷. نتیجه شبکه DCNN

ارزیابی نتایج

مقایسه میان دو نتیجه خروجی از ردیاب چشم و شبکه‌های عصبی انجام شد. در این مقایسه از سیستم شباهت‌سنجی بین تصاویر استفاده شد. در این [۴۳] پژوهش روشی برای جست و جوی تصاویر مشابه از نظر اطلاعات تصویری ارائه شده است. در این سیستم تصاویر به صورت پیکسل به پیکسل مورد مقایسه قرار می‌گیرند. دو تصویر خروجی به صورت یکسان از لحاظ اندازه بررسی می‌شوند و نور و ویژگی‌های پیکس‌های دو تصویر مورد ارزیابی قرار می‌گیرند. نمونه‌ای از امتیازات هر ستون از دو تصویر در جدول ۱ قابل مشاهده است. برای ارزیابی نتایج در این پژوهش مقایسه بین دو نتیجه ردیاب چشم و ۵ شبکه عصبی انجام شده است. جدول ۲ نتایج شباهت ۴ شبکه عصبی را برای ۳ بازی ذکر شده نشان می‌دهد. جدول ۳ نتایج ۳ معیار متفاوت را برای ۳ بازی مشخص کرده است.

جدول ۱. نمونه‌ای از نحوه مقایسه دو تصویر از لحاظ شباهت

1111111111111111001000000100010010000	1111111111111111011100000110010011111
0000000011111000011110001111000001111	0100110111110000011100000110000001111
1111111000000111111111100000001110001	010000000000000000000000000000000000
110000000111111110000000111111111000	000000000000000000000000000000000000
0001111000111000000101100001100000011	000000000000000000000001000000000000001
1100001000000011000000000010000000000	0000000000000000000000000000000011000001100
000000000000000000000000000000000000	000001000000000000000000000000000000000

جدول ۲ نتایج مقایسه ۴ شبکه عصبی

	GBVS	Itti & Koch	Grad-CAM	DCNN
Mario-Bros	۷۶,۲۶	۸۴,۶۳	۴۵,۹۳	۴۹,۰۶
Pokémon	۹۱,۸۷	۸۷,۸۹	۶۲,۵۰	۵۹,۶۹
NFS(fps)	۷۹,۸۵	۷۶	۷۰,۳۱	۷۲,۵۶
NFS(tps)	۸۸,۲۰	۹۲,۱۲	۶۸,۰۱	۶۵,۲۵

جدول ۳ نتایج ۳ معیار مختلف

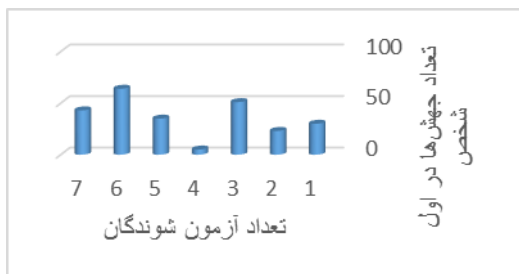
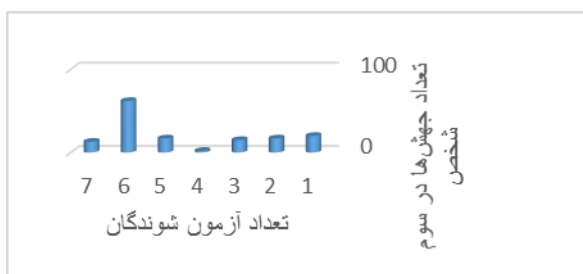
	AUC	CC	NSS
Mario-Bros	۴۸,۲۵	۷۵,۲۶	۸۲,۹۱
Pokémon	۵۷,۴۲	۷۸,۵۲	۷۹,۴۱
NFS(fps)	۷۶,۸۶	۸۶	۶۹,۱۳
NFS(tps)	۶۲	۹۱,۱۰	۸۵,۵۸

قسمت دوم از نتایج ارزیابی نتایج جهش چشم در بازی‌ها است. پیش‌تر دلیل بررسی جهش افزایش تمرکز و رابطه‌ی آن با خستگی بازیکن در بازی ذکر شد. جهش‌ها همانند نتایج قبل در ۳ بازی با دو حالت اول شخص و سوم شخص ثبت شده‌اند. تصویر ۸ نمونه‌ای نحوه ثبت شدن این نقاط در بازی NFS را نشان می‌دهد.



تصویر ۸. آزمون بازی NFS به دو صورت اول شخص و سوم شخص

نمودارهای ۳ نتایج به‌دست‌آمده از جهش‌های چشم برای بازی NFS را مشخص کرده است.



نمودار ۳. تعداد جهش‌های چشم در دو حالت اول شخص و سوم شخص

سطح معناداری جهش چشم برای ۳ بازی اندازه‌گیری شد که نتیجه ۰,۲۶ بود. این نتیجه رابطه معناداری بین دو حالت اول شخص و سوم شخص را نشان می‌دهد. با توضیحات و نتایج به دست آمده می‌توان گفت تعداد جهش چشم در حالت اول شخص همچون نگاه به یک تابلو ثابت دارای حس جست و جوی بیشتر است. نتایج جهش چشم برای ۷ آزمون شونده و سه بازی متفاوت در جدول ۴ آمده است.

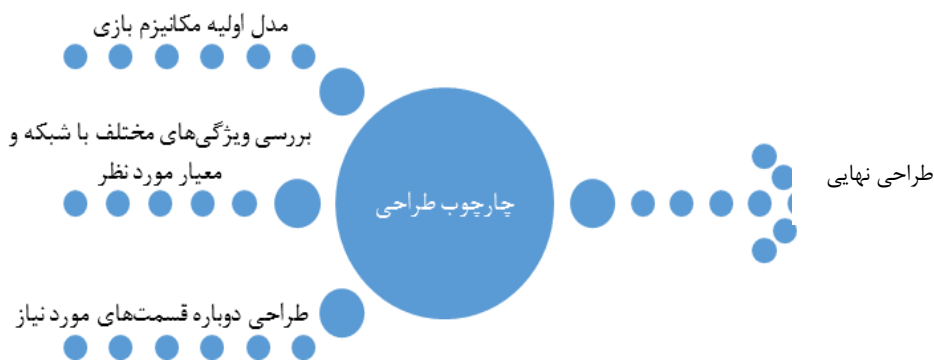
جدول ۴. نتایج جهش چشم

خطای معیار میانگین	انحراف معیار	تعداد	میانگین	
۷,۲۷۸۲۴	۱۹,۲۵۶۴۲	۷	۳۵,۸۵۷۱	گروه ۱ اول شخص
۷,۱۹۵۹۹	۱۹,۰۳۸۸۱	۷	۲۰,۸۵۷۱	سوم شخص

سطح معنی‌داری لوین	همبستگی	تعداد	
.051	.751	۷	گروه ۱ اول شخص سوم شخص

سطح معنی‌داری	درجه آزادی	t	تفاوت‌های جفتی				میانگین	انحراف معیار	خطای معیار میانگین	%۹۵ فاصله اطمینان از تفاوت
			پایین‌ترین	بالاترین	میانگین	خطای معیار				
.026	۶	۲,۹۳۹	۲۷,۴۸۸۲۷	۲,۵۱۱۷۳	۵,۱۰۳۶۹	۱۳,۵۰۳۰۹	۱۵,۰۰۰۰۰	گروه ۱ اول‌شخص و سوم‌شخص		

نتایج مقایسه ردیاب چشم با شبکه‌های عصبی کانولوشنی نشان می‌دهد بهترین نتیجه برای از لحاظ مشابهت برای بازی سه‌بعدی و حالت سوم شخص بازی NFS به دست آمده است. با این حال نتایج کلی در شبکه GBVS و Itti & Koch برای بیشتر حالات خوب پیش‌بینی شده است. استفاده از دو حالت سه‌بعدی و دوبعدی و همچنین اول‌شخص و سوم‌شخص در این پژوهش و همزمان مقایسه با چندین شبکه و معیار متفاوت چند انتخاب برای هر کدام از این مسیرها در اختیار طراحان قرار می‌دهد. در ادامه نتیجه جهش چشم و ارتباط مستقیم آن با اول‌شخص بودن بازی به کنش بیشتر بازیکن اشاره دارد که می‌توان در قسمت طراحی مکانیزم بازی با در نظر گرفتن هدف بازی آن را طراحی نمود. حالت سه‌بعدی بازی ارتباط بهتری با ردیاب چشم برقرار کرد چنان‌که بازیکن خود را در آن وضعیت واقعی حس می‌کند. حالت اول‌شخص نیز شباهت بیشتری به برخورد مستقیم بازیکن و مخاطب با یک تصویر دوبعدی در نمایشگر دارد بنابراین نتایج برای این حالت نیز در اکثر داده‌های ردیاب چشم میزان درستی بیشتری داشتند. در ادامه در نمودار ۴ پروتکل کلی و نحوه استفاده از روش پیشنهادی برای طراحی بازی پیاده‌سازی شده است.



نمودار ۴. پروتکل کلی روش پیشنهادی

قسمت‌های مختلفی از بازی ماریو در مدل پیش‌بینی GBVS مورد بررسی توصیف قرار گرفته و نحوه تحلیل بصری آن توسط طراح شرح داده شده است. شکل ۹ تعدادی از نتایج را نشان می‌دهد.



تصویر ۹ نتایج ماریو بروس از شبکه GBVS

نتایج شکل ۹ به ترتیب از سمت چپ اولین شکل نقطه خیرگی را در این صحنه به لوله سبز رنگ اختصاص داده است که با وجود اینکه جزو عناصر کلیدی در این بازی است همین طور داشتن نیم نگاهی به دشمن (سوسک قهوه‌ای) در این صحنه لازم است و همچنین نگاه گذرایی که بر روی ابر قرار دارد بنابر نداشتن کاربرد آن در بازی در اکثر صحنه‌ها لازم بنظر می‌آید

اما دلیل بر خالی کردن آسمان نیست. در تصاویر بعدی اهمیت امتیازها در بازی قابل مشاهده است که با رنگ طلایی در زمینه آبی کاملاً مشخص است. رنگی آبی و طلایی در دایره رنگی ایتن تقریباً مکمل یکدیگر هستند که رنگ طلایی امتیازها نشان دهنده امید و هیجان برای کسب امتیاز است. بازی ماریو بروس در عین سادگی دارای ساختار و طراحی بهینه و درست که مدت‌ها در ذهن مخاطبان جای گرفته است.

نتیجه‌گیری

شبکه‌های عصبی کانولوشنی نتایج بسیار خوبی در مقایسه با داده‌های استخراج شده از ردیاب چشم به دست آوردند. در دهه اخیر شاهد استفاده از سیستم‌های هوشمند مبتنی بر شبکه‌های عصبی هستیم که در پردازش تصویر بسیار قدرتمند عمل می‌کنند. در زمینه پیش‌بینی رفتارهای چشم نیز که بر پایه‌های داده‌های تصویری انجام می‌گیرد شبکه‌های عصبی با پیشرفت سخت‌افزارها نتایج بهتری می‌گیرند. پیش‌بینی رفتارهای چشم که در این پژوهش از آن بهره گرفته شد می‌تواند در بسیاری از طراحی‌های هنری و صنعتی استفاده شود. همان‌طور که در پژوهش [۴۴] برای به دست آوردن اهمیت تبلیغات برای جاذبه توریستی توسط ردیاب چشم انجام شده است. با توجه به رابطه مستقیم بازی‌های ویدیویی با چشم، ردیاب چشم تکنولوژی بسیار کاربردی در این زمینه محسوب می‌شود. با این وجود این تکنولوژی در دست تمام طراحان نیست و استفاده از آن و در اختیار داشتن تمام قشر بازیکنان برای گرفتن آزمون دور از دسترس به نظر می‌رسد. بنابراین ایجاد مدل شبکه عصبی کانولوشنی با داده‌های انواع بازی در نتایج پیش‌بینی رفتارهای چشم ضروری به نظر می‌رسد. از طرفی ابعاد بازی‌ها در حال افزایش است و تعیین محدودیت برای آن کمی دشوار به نظر می‌رسد. تقسیم‌بندی به اول‌شخص و سوم شخص برای بازی‌ها باعث به وجود آمدن دو حالت کلی شد که نتیجه آن تفاوت در نحوه نگرش بازیکن برای یک بازی در این دو حالت بود. نتایج بازی‌های سه‌بعدی به دلیل نزدیک بودن به حالت طبیعی نگاه مخاطب با دقت بیشتری حاصل شد که این از دیگر نتایج مفید برای طراحان به حساب می‌آید. طراحان بازی با استفاده از شبکه‌های عصبی و نتایج پیش‌بینی وابسته به نوع بازی به آنالیز بصری بازی پی برده و تمهیدات لازم را برای ادامه یا تکمیل بازی بررسی می‌کنند. نتایج بررسی باعث افزایش آگاهی طراح به تمام قسمت‌ها و نحوه ارتباط آن‌ها با یکدیگر و مکانیک بازی می‌شود. از سختی‌های این نوع تحلیل برای طراحان استفاده همزمان از اکثر شبکه‌های پیشنهادی است. طراحی یک شبکه برای اکثر نقاط تحلیل در بازی از پژوهش‌های آینده خواهد بود که در سیستمی ساده و قابل استفاده در دسترس طراحان بازی قرار گیرد. در فرایند انتخاب بازی‌های پژوهش سعی بر آن شد تا از بازی‌های محبوب و آشنا برای آزمون استفاده شود. بازی ماریو از بازی‌های پرتعدادی است که هنوز نسخه‌های جدیدی از آن رونمایی می‌شود و می‌تواند الگویی برای طراحان بازی باشد.

منابع و مراجع

- [1] Nelson, M.J. and M. Mateas(2007) Towards automated game design. in Congress of the Italian Association for Artificial Intelligence. Springer.
- [2] Cook, D (2007) The chemistry of game design. World Wide Web electronic publication.,
- [3] Theodosiou, S. and I. Karasavvidis (2015) Serious games design: A mapping of the problems novice game designers experience in designing games. Journal of e-Learning and Knowledge Society,. 11(3.)
- [4] Björk, S. and J. Holopainen (2006) Games and design patterns. The game design reader: p. 410-437.
- [5] Aleven, V., et al (2010) Toward a framework for the analysis and design of educational games. in 2010 third IEEE international conference on digital game and intelligent toy enhanced learning. IEEE.
- [6] Renshaw, T., R. Stevens, and P.D. Denton (2009) Towards understanding engagement in games: an eye-tracking study. On the Horizon.,
- [7] Schrom-Feiertag, H., V. Settgest, and S. Seer (2017) Evaluation of indoor guidance systems using eye tracking in an immersive virtual environment. Spatial Cognition & Computation. 17(1-2): p. 163-183.
- [۸] مصطفوی، س. و. ر. س. ی. س. و. ب. د. ص. (۱۳۹۶) ارزیابی روشی در جهت بهبود روند طراحی مرحله در بازی های رایانه ای با استفاده از دستگاه ردیاب چشم. سومین کنفرانس ملی بازی های رایانه ای؛ فرصت ها و چالش ها، اصفهان، دانشگاه اصفهان.
- [9] Cornia, M., et al(2018) Predicting human eye fixations via an lstm-based saliency attentive model. IEEE Transactions on Image Processing. 27(10): p. 5142-5154.
- [10] Vidal, M., et al (2012) Wearable eye tracking for mental health monitoring. Computer Communications. 35(11): p. 1306-1311.
- [11] Ettinger, U., et al (2003) Reliability of smooth pursuit, fixation, and saccadic eye movements. Psychophysiology. 40(4): p. 620-628.
- [12] Arolt, V., et al (1998) Distinguishing schizophrenic patients from healthy controls by quantitative measurement of eye movement parameters. Biological psychiatry. 44(6): p. 448-458.
- [13] Pons, J., et al (2017) Timbre analysis of music audio signals with convolutional neural networks. in 2017 25th European Signal Processing Conference (EUSIPCO). IEEE.
- [14] Takahashi, R., et al (2018) A system for three-dimensional gaze fixation analysis using eye tracking glasses. 5(4): p. 449-457.
- [15] Salehin, M.M. and M. Paul (2017) A novel framework for video summarization based on smooth pursuit information from eye tracker data. in Multimedia & Expo Workshops (ICMEW), 2017 IEEE International Conference on. IEEE.
- [16] Li, B., B. Mettler, and J. Andersh (2015) Classification of human gaze in spatial guidance and control. in Systems, Man, and Cybernetics (SMC), 2015 IEEE International Conference on. IEEE.
- [17] Zhang, X. and S.-M.J.I.A. Yuan (2018) An Eye Tracking Analysis for Video Advertising: Relationship Between Advertisement Elements and Effectiveness. 6: p. 10699-10707.
- [18] Tatler, B.W (2007) The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions. Journal of vision. 7(14): p. 4-4.

- [19] Roy, A.K., et al (2017) A Novel Technique to develop Cognitive Models for Ambiguous Image Identification using Eye Tracker.(1): p. 1-1.
- [20] Yin, N. and M.G.J.J.o.H.S.N. Hluchyj (1993) Analysis of the leaky bucket algorithm for on-off data sources. 2(1): p. 81-98.
- [21] Majaranta, P. and A. Bulling (2014) Eye tracking and eye-based human-computer interaction, in *Advances in physiological computing*. Springer. p. 39-65.
- [22] Mallick, R., et al (2016) The use of eye metrics to index cognitive workload in video games. in *2016 IEEE Second Workshop on Eye Tracking and Visualization (ETVIS)*. IEEE.
- [23] Rudoy, D., et al (2013) Learning video saliency from human gaze using candidate selection. in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.
- [24] Krejtz, K., et al (2014) Entropy-based statistical analysis of eye movement transitions. in *Proceedings of the Symposium on Eye Tracking Research and Applications*. ACM.
- [25] Tien, T., et al (2015) Differences in gaze behaviour of expert and junior surgeons performing open inguinal hernia repair. 29(2): p. 405-413.
- [26] Sarter, N.B., R.J. Mumaw, and C.D.J.H.f. Wickens (2007) Pilots' monitoring strategies and performance on automated flight decks: An empirical study combining behavioral and eye-tracking data. 49(3): p. 347-357.
- [27] Frutos-Pascual, M. and B. Garcia-Zapirain (2015) Assessing visual attention using eye tracking sensors in intelligent cognitive therapies based on serious games. *Sensors*. 15(5): p. 11092-11117.
- [28] Almeida, S., Ó. Mealha, and A. Veloso (2016) Video game scenery analysis with eye tracking. *Entertainment Computing*. 14: p. 1-13.
- [29] Polonio, L., S. Di Guida, and G. Coricelli (2015) Strategic sophistication and attention in games: An eye-tracking study. *Games and Economic Behavior*. 94: p. 80-96.
- [30] Devetag, G., S. Di Guida, and L. Polonio (2016) An eye-tracking study of feature-based choice in one-shot games. *Experimental Economics*. 19(1): p. 177-201.
- [31] Wang, L., et al (2015) Deep networks for saliency detection via local estimation and global search. in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.
- [32] Kümmerer, M., T.S. Wallis, and M.J.a.p.a. Bethge (2016) DeepGaze II: Reading fixations from deep features trained on object recognition .
- [33] Xu, J., et al (2014) Predicting human gaze beyond pixels. 2014. 14(1): p. 28-28.
- [34] Kummerer, M., T.S. Wallis, and M. Bethge (2018) Saliency benchmarking made easy: Separating models, maps and metrics. in *Proceedings of the European Conference on Computer Vision (ECCV)*.
- [35] Wloka, C., I. Kotseruba, and J.K. Tsotsos (2018) Active fixation control to predict saccade sequences. in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.
- [36] Selvaraju, R.R., et al (2019) Grad-cam: Visual explanations from deep networks via gradient-based localization. in *Computer Vision and Pattern Recognition (cs.CV); Artificial Intelligence (cs.AI); Machine Learning (cs.LG)*.
- [37] Harel, J., C. Koch, and P. Perona (2007) Graph-based visual saliency. in *Advances in neural information processing systems* .
- [38] Itti, L., C. Koch, and E. Niebur (1998) A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on pattern analysis and machine intelligence*. 20(11): p. 1254-1259.

- [39] Itti, L. and C. Koch (2000) A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision research*. 40(10-12): p. 1489-1506.
- [40] Itti, L. and P.F. Baldi (2006) Bayesian surprise attracts human attention. in *Advances in neural information processing systems* .
- [41] Selvaraju, R.R., et al (2017) Grad-cam: Visual explanations from deep networks via gradient-based localization. in *Proceedings of the IEEE international conference on computer vision* .
- [42] Simonyan, K., A. Vedaldi, and A. Zisserman (2013) Deep inside convolutional networks: Visualising image classification models and saliency maps. arXiv preprint arXiv:1312.6034.
- [43] Suryawanshi, D (2018) Image Recognition: Detection of nearly duplicate images, California State University Channel Islands.
- [44] Li, Q., Z.J. Huang, and K. Christianson (2016) Visual attention toward tourism photographs with text: An eye-tracking study. *Tourism Management*. 54: p. 243-258.