ششمین کنفرانس بین‌المللی
«بازی‌های رایانه‌ای؛ فرصت‌ها و چالش‌ها»
۳۰ بهمن و ۱ اسفند ۱۳۹۹ – دانشگاه اصفهان

مروری بر روش‌های پردازش تصویر با تاکید بر شبکه های مولد رقابتی و کاربرد کارتونی‌سازی در فضای بازی‌های رایانه‌ای

راحله حیاتی*[۱]، بهنام علیزاده اشرفی[۲]، حمیده حرمتی[۳]، باقر بهرام شتربان[۴] ، محمد ببردل بناب[۵]

۱- **دانشجوی کارشناسی ارشد دانشگاه هنراسلامی تبریز،دانشکده چندرسانه‌ای**
Email: rahil.hayati14@gmail.com
۲- **استادیار دانشگاه هنر اسلامی تبریز، دانشکده چندرسانه‌ای**
Email: b.alizadehashrafi@tabriziau.ac.ir
۳- **استادیار دانشگاه هنر اسلامی تبریز، دانشکده چندرسانه‌ای**
Email: hormati@tabriziau.ac.ir
۴- **مربی دانشگاه هنر اسلامی تبریز، دانشکده چندرسانه‌ای**
Email: b.bahram@tabriziau.ac.ir
۵- **مربی دانشگاه هنر اسلامی تبریز، دانشکده چندرسانه‌ای**
Email: babrdel@utar.edu.my

چکیده

این مقاله با هدف انتشار نتایج کارتونی‌سازی به فضای بازی ارائه شده‌است و براساس مطالب بخش پیشینه پژوهش پایان‌نامه کارشناسی ارشد با عنوان «تبدیل کاراکتر ویدیو به کاراکترمجازی‌متحرک در زمان‌واقعی با استفاده از پردازش‌تصویر»، ارائه شده‌است. در واقع، این مقاله در مورد انتشار یکی از کاربردهای نتایج کارتونی‌سازی ارائه شده‌است. منظور از کارتونی‌سازی، تبدیل تصاویر واقعی به تصاویر کارتونی از طریق کامپیوتر است. به همین منظور، به مرور در مورد انیمیشن‌سازی و الگوریتم‌های پردازش‌تصویر بدون استفاده از شبکه عصبی و با استفاده از شبکه عصبی(بهینه‌سازی تصویر،Feed-Forward، GAN) پرداخته‌است. در این مقاله با استفاده از نتایج آموزش شبکه عصبی عمیق تحت عنوان Neural Style Transfer، یک پروتوتایپ برای بررسی کارتونی‌سازی فضای بازی ارائه شده‌است و در بخش نتیجه‌گیری، تاثیر کارتونی‌سازی بازی برای جذابیت بازی، تجربه‌ی جدید حس حضور در فضای کارتونی، حس خیال و انتزاع مطرح شده‌است.

کلمات کلیدی: شبکه عصبی عمیق، انتقال‌سبک عصبی، Feed-Forward & GAN، کارتونی‌سازی،بازی‌های رایانه‌ای

۱–مقدمه

بیش از ۹۰ درصد اطلاعات اطرافمان توسط مشاهده دریافت می‌شود، حس بینایی سرعت عمل بسیاری دارد(اصلانی، مرادپور، و نهاردانی ۱۳۹۱، ۱۲)، به همین علت زبان بصری تا حدود بسیار زیادی بر انسان تاثیرگذار است. زبان بصری یک زبان همگانی است که مشترک بین تمامی انسان‌ها است و به یک ملیت یا قومیت خاص، اختصاص ندارد، به همین علت، نقاشی‌هایی ترسیم شده توسط انسان اولیه بر روی دیواره‌ی غارها، به سادگی پیام انسان‌های اولیه را به همه‌ی انسان‌ها انتقال می‌دهد. این تصاویر با اینکه بصورت ساده طراحی شده‌اند، نشان‌دهنده‌ی حس جستجو، حرکت، زندگی، شکار و خیلی از موارد است. اگر بخوبی تصاویر و نشانه‌های غارنوشته‌ها را دنبال کنیم و خوب دقیق شویم، شاید بتوانیم به این تصاویر از نگاه بازی نگاه کنیم. همانطور

که خیلی از افسانه‌ها یا فیلم‌ها بر اساس همین تصاویر ساخته شده اند. در انیمیشن و گرافیک، نیز از ساده‌سازی عناصر بصری برای انتقال سریع مفاهیم استفاده می‌شود. انیمیشن‌های اولیه با اینکه بدون صدا ساخته شده‌اند، بسیار تاثیرگذار هستند. با این تفاسیر می‌توانیم نتیجه بگیریم که استفاده از طراحی‌های دو بعدی در فضای بازی می‌تواند تاثیر گذار باشد.

با همین دید، برای کارتونی‌سازی به شیوه اتومات به تحقیق و تفحص پرداخته شد. در بخش پیشینه، ابتدا در مورد انیمیشن و روش‌های آن مرور شد و سپس الگوریتم‌های پردازش تصویر در دو بخش(بدون استفاده از CNN و با استفاده از CNN) بررسی شد. در بخش ارائه روش، کارتونی‌سازی فضای بازی به روش پروتوتایپ ارائه شده است. برای کارتونی‌سازی فضای بازی از نتایج روش «یانگ چن و همکارانش[۱]» استفاده شده‌است. سپس به نتیجه‌گیری و بحث پرداخته شده‌است و پیشنهادات برای کارهای آتی ارائه شده است.

۲– پیشینه تحقیق

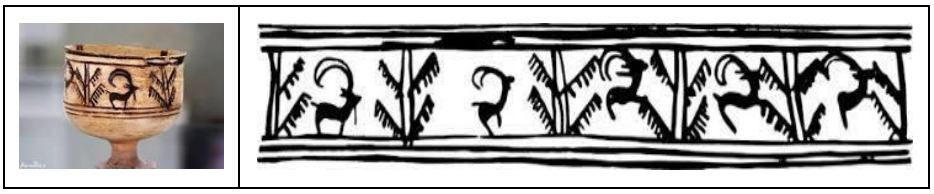
در این بخش، ابتدا، انیمیشن و روش‌های آن مورد بررسی قرار گرفته است. سپس الگوریتم‌های پردازش تصویر در دوبخش الگوریتم‌های پردازش تصویر بدون استفاده از CNN و الگوریتم‌های پردازش تصویر با استفاده از CNN بررسی شده‌است.

۱–۲– انیمیشن و روش‌های آن

در این زیربخش ابتدا در مورد تداوم دید و پدیده‌ی فای و سپس درمورد انیمیشن و انواع آن، وسایل اولیه ایجاد انیمیشن و تکنیک‌های آن توضیح ارائه شده‌است.

اصل تداوم دید و پدیده فای، از اصول اولیه درک حرکت در تماشای فیلم‌ها و انیمیشن‌ها می باشد. **اصل تداوم دید:** نظریه‌ای که در سال ۱۸۲۴ توسط یک پزشک بریتانیایی به نام «پیترمارک رژه»[1] در مورد چشم انسان بصورت علمی مطرح شد: «مغز انسان تصویری را که بر شبکیه چشمش می‌افتد تا زمانی حدود یک بیستم تا یک پنجم ثانیه، پس از کنار رفتن تصویر از مقابل چشم همچنان حفظ می‌کند»(بزرگمهر ۱۳۸۹). به همین دلیل ذهن هیچ تصوری از پلک‌زدن ندارد، زیرا درحالیکه پلک‌زدن چشم در مدت کمتر از یک ثانیه رخ می‌دهد، تجسمات پیش از پلک‌زدن، محفوظ می‌ماند(بزرگمهر ۱۳۸۹). به همین دلیل هنگام تماشای یک سری تصاویر که با سرعت معینی از هم عبور می‌کند، توهم حرکت در ذهن پدید می‌آید و آن تصاویر بصورت مداوم دیده می‌شوند. **پدیده فای:** هنگام مشاهده پره‌های یک پنکه در حال حرکت، پره‌های آن بصورت تیغه ای واحد و مدور دیده می‌شود، یا پره‌های رنگارنگ یک چرخ گردان، بصورت تک رنگی آمیخته دیده می‌شود، که به این پدیده فای گفته می‌شود که توسط «مارکس ورتایمر»[2] مورد مطالعه قرار گرفت(بزرگمهر ۱۳۸۹). «پدیده فای» به کمک «تداوم دید»، توهم حرکت را ایجاد می کند و بنابراین هنگام نمایش تصاویر ساکن به مدت ۱۲ تا ۲۴ بار در ثانیه، تصاویر همچون تصاویر متحرک مشاهده می‌شود (بزرگمهر ۱۳۸۹). **انیمیشن**، توهم حرکتی است که از پشت‌سرهم قرار گرفتن تصاویر در ذهن بوجود می‌آید. در بعضی از فیلمهای نقاشی متحرک برای هر ثانیه ۱۲ فریم نقاشی می‌کنند و به هنگام فیلمبرداری از هر نقاشی ۲ فریم فیلم می‌گیرند. تاریخچه انیمیشن به نقاشی‌های انسان های اولیه بر دیوار غارها برمی‌گردد. سفالینه قدیمی از شهر سوخته در سیستان و بلوچستان وجود دارد که روی آن نقش‌هایی بصورت متوالی دیده می‌شود که می‌توان گفت این اولین انیمیش ساخته‌شده است(شکل۱)("پویانمایی" ۱۳۹۹).

1 Peter Mark Roget
2 Marx Vertimer



شکل ۱: سفالینه قدیمی از شهر سوخته در سیستان و بلوچستان("پویانمایی" ۱۳۹۹)

بلافاصله بعد از مطرح شدن نظریه «تداوم دید»، سایر دانشمندان این نظریه را با حرکات صفحات و سکه های گردان ، یا با به سرعت ورق زدن دفترچه ای پر از تصویر و غیره، به آزمایش گذاشتند و تایید کردند ("پویانمایی" ۱۳۹۹). وسایل اولیه‌ای که برای ایجاد انیمیشن و آزمایش توهم حرکت ایجاد شدند، در جدول ۱ شرح داده شده‌است. نکته جالب اینجاست که این وسایل ابتدا بعنوان وسایل بازی مورد استفاده قرار گرفت، همچنین شاید بتوانیم بگوییم سفالینه قدیمی(شکل ۱)، نیز به قصد بازی ساخته شده بوده است.

جدول ۱: وسایل‌اولیه ایجاد انیمیشن(منبع:نگارنده)

عنوان	شرح	تصویر
توماتروپ	یک دایره مقوایی است که دوطرف آن تصویری طراحی شده‌است و هنگامیکه مقوا به وسیله نخ دو طرف آن، با انگشتان دست به سرعت چرخانده شود، نور هر دو تصویر در چشم ضبط می‌شود و یک تصویر از هر دو طرف مشاهده می‌شود. این نشان می‌دهد که چشم تصاویر را هنگامیکه بصورت متوالی در یک زمان به نمایش در می‌آید را حفظ می‌کند("توماتروپ" ۱۳۹۸).	
فنکنیستوپ	ورقی گرد که لبه‌های آن دارای شکاف‌هایی است که یک‌سری نقاشی بصورت متوالی روی آن ترسیم می‌شود. بیننده ورق را جلوی آینه نگه می دارد و از میان شکاف‌های آن نگاه می‌کند، با چرخش زیاد این ورق، تصاویر بصورت متحرک نشان داده می‌شود(Phenakistiscope" 2020 ").	
زوتروپ	گردونه‌ای استوانه‌ای شکل است که درون آن نوار کاغذی حاوی یک رشته تصاویر نقاشی شده در آن قرار می‌گیرد و از چرخش حرکت گردونه توهم حرکت ایجاد می‌شوند(Zoetrope" 2020 ").	
فلیپ‌بوک	دفترچه‌ای که روی تمام صفحات آن نقاشی ترسیم شده و با ورق زدن سریع آن تصاویر بصورت متحرک مشاهده می‌شوند(Flip book"2020").	

انیمیشن تنها نمایش داستان فیلم‌های کارتونی نیست، انیمیشن زمینه‌های بسیاری را دربرمی‌گیرد، از جمله: در بازی‌های ویدیویی، رابط‌های کاربری، فضای مجازی، واقعیت افزوده، بازی‌های کامپیوتری، گرافیک‌های حرکتی و جلوه‌های بصری. بطور کلی انیمیشن‌های رایانه‌ای به دو دسته‌ی انیمیشن‌های دو بعدی و سه‌بعدی تقسیم می‌شوند(Animation"2020"). در جدول ۲، تکنیک‌های ساخت انیمیشن شرح داده شده‌است.

جدول ۲: تکنیک‌های ساخت انیمیشن(منبع:نگارنده)

نام تکنیک	شرح
انیمیشن روی طلق	دراین روش ابتدا طرح روی طلق نقاشی و رنگ‌آمیزی می‌شود و سپس طلق‌ها را روی تصویر پس‌زمینه قرار می‌دهند و دو فریم از هر تصویر فیلمبرداری می‌شود(سارانی‌نژاد ۱۳۹۱).
انیمیشن روی کاغذ	طراحی هر صحنه روی کاغذ انجام می‌شود و فیلمبرداری می شود. (سارانی‌نژاد ۱۳۹۱).
انیمیشن بدون دوربین	از ظاهر کردن نوارفیلم خام و سپس طراحی با مدادچرب، مرکب یا رنگ روغن یا وسایل دیگر روی نوار فیلم ایجاد می‌شود، یا با ایجاد خراش‌هایی بر روی نوار فیلم ایجاد می‌شود. (سارانی‌نژاد ۱۳۹۱).
انیمشن کات اوت یا بریده مقوا	یعنی یا استفاده از موادی مثل مقوا، کاموا، کاغذهای مختلف یا مواد دیگر بصورت کولاژ به حرکت درآورده می‌شود و فیلمبرداری می‌شود، یا حتی بصورت نوردهی از پشت این مقوا، تصاویر سایه‌واری را ایجاد می‌کنند(سارانی‌نژاد ۱۳۹۱).
انیمیشن سه‌بعدی	عروسک‌هایی را که حالت مفصل‌های فرم پذیر دارند می‌سازند یا از حجم‌ها، اجسام یا خمیرهایی استفاده می‌شود و حرکات فیلمبرداری می‌شود. (سارانی‌نژاد ۱۳۹۱).
کلیمیشن	در این روش پرسوناژها وپس‌زمینه با گل ساخته می‌شد و سپس پرسوناژها حرکت داده می‌شد. امروزه از مواد مشابه مثل خمیر مجسمه‌سازی، موم یا موارد دیگر استفاده می‌شود(سارانی‌نژاد ۱۳۹۱).
انیمیشن با عکس	عکس‌ها روی پس‌زمینه کلاژ می‌شوند و یا تغییراتی روی این عکس‌ها صورت می‌گیرد و فیلمبرداری می‌شود. (سارانی‌نژاد ۱۳۹۱).
انیمیشن پیکسلیشن	حرکات بازیگر انتخاب می‌شود و با تغییر سرعت این حرکات، جلوه‌های حرکتی ایجاد می‌شود. (سارانی‌نژاد ۱۳۹۱)
انیمیشن تخته سنجاق	با چیدن یک سری سوزن و نورپردازی و حرکت دادن این سوزن‌ها، جلوه‌های حرکتی ایجاد می‌شود(سارانی‌نژاد ۱۳۹۱).
انیمیشن زیر دوربین	مواد ریزی مثل شن یا خرده‌های چوب یا مواد ریز دیگری را بر روی میز نور، قرار می‌دهند و طرح هایی را بر روی آن ایجاد می کنند و فیلمبرداری می‌شود. (سارانی‌نژاد ۱۳۹۱).
دایینامیشن	از ترکیب صحنه‌های عروسکی با صحنه‌های دنیای واقعی ایجاد می‌شود. (سارانی‌نژاد ۱۳۹۱).
دیاگرام انیمیشن	در این روش از شکل‌های ساده، خطوط و فلش برای نشان دادن موضوعات ساده استفاده می‌شود و بیشتر جنبه‌ی آموزشی دارد. (سارانی‌نژاد ۱۳۹۱).
انیمیشن تخته سفید	در این روش داستان روی پس‌زمینه سفید طراحی می‌شود. معمولا نقاشی‌ها بصورت سیاه رنگ طراحی می‌شود و همچنین از تصویر واقعی دست انیماتور، استفاده می‌شود. در این روش معمولا بطور مداوم بصورت ساده و جذاب طراحی می‌شود و از جلوه‌های تصویری بیش از حد، اجتناب می‌شود تا بر پیام اصلی داستان تاکید شود. معمولا از این نوع انیمیشن برای محتواهای آموزشی استفاده می‌شود(Whiteboard animation"2020")
استاپ موشن	در این روش از یک یا چند شی عکس گرفته می‌شود و بعد مکان اشیا را کمی جابجا می‌کنند و عکس گرفته می‌شود و این فرآیند جابجایی آنقدر تکرار می‌شود تا بعد از نمایش پشت سرهم این عکس‌ها توهم حرکت اتفاق می‌افتد تا مفهوم اصلی محتوا را برساند.(Animation Styles & Types"2020").
اینفوگرافیک	برای نمایش منظم مفاهیم پیچیده و جزئیات زیاد، به روش بسیار ساده و در زمان بسیار کوتاه استفاده می‌شود. این روش شبیه به حرکت در آوردن چند اسلاید ساده است. .(Animation Styles & Types"2020").
ایزومتریک	همانند اینفوگرافیک است با این‌تفاوت که از پیکتوگرام‌های طراحی شده موجود استفاده می‌شود تا محتوا را برساند. .(Animation Styles & Types"2020").
مت۱	در این روش بخشی از لنز دوربین پوشیده می‌شود و فیلمبرداری می‌شود و دوباره قسمت پوشیده‌شده از لنز برداشته می‌شود و فیلمبرداری می‌شود. با این روش جلوه‌های بصری جالبی بصورت ترکیبی ایجاد می‌شود[۲].

[1] Matte

روتسکوپی	ابتدا از حرکت بازیگران، فیلمبرداری می‌شود و سپس طرح‌های جدیدی از تصاویر حرکت بازیگران، طراحی می‌شود و فیلمبرداری می‌شود که این شیوه در بیشتر انیمیشن‌های «والت دیزنی» استفاده شده‌است[۲].

۲-۲- تعریف سبک و تفاوت آن با بافت

سبک، «سبک در لغت به معنای طور و عمل و طرز و روش و قاعده و شیوه است("سبک" ۱۳۹۹)». سبک در واقع شیوه‌ی هنری هنرمند در اجرای اثر هنری است. تفاوت سبک و بافت، «سبک با بافت بسیار مرتبط است اما محدود به بافت نیست. سبک همچنین میزان زیادی از ساده‌سازی و جلوه‌های انتزاعی شکل را در برمی گیرد، که به ترکیب‌بندی یا ترازبندی ویژگی‌های سبک بر می‌گردد[۳]».

۳-۲- تعریف یادگیری ماشین و شبکه عصبی کانولوشنال

اولین‌بار «آرتور ساموئل»، واژه‌ی **یادگیری ماشین**، را ثبت کرد، یادگیری ماشین شاخه‌ای از علوم کامپیوتر است که بدون انجام برنامه‌نویسی صریح، به کامپیوتر توانایی یادگیری می‌بخشد(ابراهیم ۱۳۹۶)»، یادگیری ماشین تکنیکی است که اجازه می‌دهد تا سیستم کامپیوتری با تجربه و داده‌ها بهبود یابد. عملکرد این یادگیری‌عمیق به کامپیوتر اجازه می‌دهد مفاهیم پیچیده را از مفاهیم ساده‌تر بسازد. درواقع یادگیری ماشین رویکرد قابل اعتمادی برای ساخت سیستم‌های هوش‌مصنوعی است که می‌تواند در محیط پیچیده و دنیای واقعی کار کند [۴]. **شبکه عصبی کانولوشنال(CNN):** شبکه‌های عصبی هستند که در آن، از عملیات ریاضی کانولوشن استفاده می‌شود((Goodfellow, Bengio, and Aaron Courville (2017).

۴-۲- پردازش تصویر

پردازش‌تصویر دیجیتال، شاخه‌ای از علوم رایانه است و با پردازش سیگنال دیجیتال سروکار دارد. سیگنال دیجیتال، نماینده تصاویر برداشته شده توسط دوربین دیجیتال یا اسکن‌شده توسط اسکنر هستند. درواقع با انجام فرآیندهایی بر روی تصاویر دیجیتال، خروجی مشخص و معینی از آن دریافت می‌شود(باختری، بیرجن ۱۳۹۸). «پردازش تصویر در معنای خاص عبارت است از هرنوع پردازش سیگنال که ورودی آن یک تصویر است مثل عکس یا صحنه‌ای از فیلم»(شیبانی صفت ۱۳۹۵). بطور کلی پردازش تصویر به دو بخش بهبود تصاویر و بینایی‌ماشین تقسیم می‌شود. در بینایی‌ماشین از الگوریتم‌هایی استفاده می‌شود که کامپیوتر بتواند توسط دوربین محیط اطراف خود را درک کند و تصمیم‌گیری کند. هدف از بهبود تصاویر، تبدیل تصویر برای یک کاربرد خاص است طوری که تصویر حاصل بهتر از تصویر اولیه باشد (باختری، بیرجن ۱۳۹۸؛ شیبانی صفت ۱۳۹۵).

محققین جامعه گرافیک کامپیوتری، الگوریتم‌های کامپیوتری را در دو بخش«اجرای فتورئالیستی» و اجرای«غیر فتورئالیستی» ارائه دادند. محققین، دربخش«اجرای غیرفتورئالیستی»، الگوریتم‌هایی را تحت عنوان«تبدیل اتومات تصویر به کارهای هنری» ارائه دادند[۳]. موضوع کارتونی‌سازی تصویر از نوع «اجرای غیرفتورئالیستی» است که یکی از موضوعات داغ محققین در تحقیقات چند رسانه‌ای، برای تبدیل تصاویر دنیای واقعی به تصاویر کارتونی است و مطالعات در این مورد، به سرعت گسترش یافته‌است و درحال گسترش است.

۱-۴-۲- روش‌های پردازش تصویر بدون استفاده از CNN

الگوریتم‌های اجرای هنری مبتنی بر تصویر بدون استفاده از CNN، طبق طبقه بندی«کیپریانیدیس و همکارانش» در جدول ۲ شرح داده شده‌است[۵] و در ادامه، سپس چند نمونه از این الگوریتم‌ها توضیح داده شد.

جدول ۳: روش‌های اجرای هنری مبتنی بر تصویر بدون CNN توسط «کیپریانیدیس و همکارانش»[۵]، (منبع: نگارنده)

روش	شرح
اجرای مبتنی بر استروگ	در این روش استروگ‌های مجازی به تدریج بر روی تصویر دیجیتال قرار داده می‌شوند تا اینکه عکس شبیه به یک سبک هنری بنظر برسد. استروگ‌هایی مثل قلم‌مو، نقطه،کاشی و غیره. در این روش تابع هدف تعریف می‌شود تا بصورت تکراری جایگذاری استروک‌ها را هدایت کند[۵]. مزیت: در شبیه‌سازی یک سبک خاص مثل نقاشی رنگ و روغن، آبرنگ و طراحی موثرهستند. محدودیت: عدم توانایی در شبیه‌سازی سبک دلخواه[۳, ۵].
تکنیک‌های مبتنی بر ناحیه	در این روش الگوهای استروگ مختلفی در نواحی خاص از تصویر قرار داده می‌شوند. یک روشی هم ارائه شده که هندسه نواحی خاصی از تصویر را تغییر می‌دهد، در این روش نواحی خاصی از تصویر به شکل‌های ابتدایی تغییر می‌یابند[۵]. محدودیت: عدم توانایی در انتقال سبک دلخواه[۳, ۵].
اجرای مبتنی بر مثال (آنالوگ‌های تصویر)	در این روش تابع هدف، انتقال آنالوگ از یک مجموعه از زوج تصاویر(سبک دار و بدون سبک) را در حالت نظارت شده یاد می‌گیرد[۵]. مزیت: برای شبیه‌سازی سبک‌های مختلف موثر است. محدودیت: زوج داده‌های تصویر معمولا دردسترس نیستند، آنالوگ‌های تصویر فقط از ویژگی‌های سطح پایین تصویر استفاده می‌کنند و در ضبط موثر محتوا و سبک شکست می‌خورند[۳, ۵].
پردازش تصویر و فیلترینگ(اجرای مبتنی بر فیلترکردن تصویر)	در این روش فیلترهای مختلف پردازش تصویر(فیلترهای دوطرفه و فیلترهای مختلف گوسین) ترکیب می‌شوند تا سیستم عکس را همانند تصاویرهنری (در حالت ساده‌سازی و انتزاع) اجرا بگیرد[۵]. مزیت: الگوریتم‌های مبتنی بر فیلتر کردن تصویر به پیاده‌سازی و کارایی در عمل روبه جلو انتشار می‌یابند. محدودیت: در تنوع سبک خیلی محدود هستند[۳, ۵].

«**ایان باک و همکارانش**» در سال ۲۰۰۶، یک الگوریتم برای ایجاد انیمیشن در زمان‌واقعی ارائه دادند که در آن، آثارهنری دستی یا نقاشی‌ها، عبارات و حرکات صورت انسان(سخنران) را تقلید می‌کند. این یک الگوریتم درون‌یابی بیان چهره است که داده‌های ردیابی شده را روی فضای دوبعدی نقاشی می‌اندازد و از تکنیک Delaunay triangulation برای پیداکردن شبیه‌ترین هندسه چهره و محاسبه وزن آن‌ها استفاده می‌کند [۶]. «**جان اریک کیپریانیدیس و همکارش**» در سال ۲۰۰۸، یک چهارچوب پردازش‌تصویر «غیر فتورئالیستیک» برای انتزاعی‌سازی تصویر ارائه دادند که تصاویر و ویدیوها و تفاسیر سه‌بعدی را سبک‌دار می‌کرد. یک فیلترینگ دوجهته که بصورت تکراری در جهت گرادیانت کار می‌کند و در نواحی لبه‌های تصویر با کنتراست پایین، لبه‌ها را ذخیره می‌کند. در این روش فیلترینگ دوجهته در جهت گرادیانت، لبه‌های میانی را در جهت عمود تغییر می‌دهد تا در لبه‌های منحنی یک خروجی صاف را تولید کند. همچنین برای لبه‌های برجسته نیز یک فیلتر گوسین یک‌بعدی در جهت گرادیانت اعمال کردند. نمونه نتایج کارشان در شکل نمایش داده شده‌است[۷]. **تکنیک cell shading** ، تکنیکی است که در آن اشکال سه‌بعدی به شکل سایه ساده رندر می‌گیرند و از آن برای ایجاد کارتون و ساخت بازی و ویدیو بهره می‌برند[۸].

۲-۴-۲- روش‌های پردازش تصویر با استفاده از CNN

در جامعه دیدکامپیوتری، مسئله‌ی تبدیل تصاویر به تصاویر هنری، معمولا بعنوان یک مسئله«سنتز بافت[۱]» مورد بررسی قرار گرفته‌است[۳]. با موفقیت کار گیتس و همکارانش[۹] در استفاده از شبکه عصبی برای ایجاد سبک‌های معروف نقاشی بر روی تصاویر طبیعی، موضوع انتقال‌سبک عصبی مطرح شد و تحقیقات تحت این عنوان گسترش یافته است. شبکه‌های انتقال‌سبک عصبی، شبکه‌هایی که در آن، ظاهر یک تصویر با حفظ محتوا به شکل‌هنری تغییر می‌یابد. در انتقال‌سبک«سیستم یک تصویر ورودی را دریافت می‌کند و به تصویر خروجی انتقال می‌یابد[۱۰]».

بطور کلی روش‌های انتقال‌سبک به دو روش **بهینه‌سازی تصویر و بهینه‌سازی مدل تقسیم‌بندی** می‌شوند[۳]. در روش **بهینه‌سازی تصویر**، الگوریتم شبکه عصبی بر اساس کاهش گرادیانت فضای تصویر را بهینه‌سازی می‌کند یعنی مکررا، بازسازی تصویر را انجام می‌دهد تا زمانی‌که تصویر سبک‌دار شده با حفظ ویژگی‌های محتوایی تصویر واقعی، نمایش ویژگی مشابه با تصویر هنری داشته باشد [۳]. در روش **مبتنی بر بهینه سازی مدل**، در این روش یک شبکه عصبیFeed-forward بصورت آفلاین آموزش می‌بیند و وزن‌های آن ثابت می‌شود و در مرحله تست، فرآیند سبک‌سازی تصویر، تنها با یک عبور از شبکه انجام می‌شود، در این روش بار محاسباتی در مرحله آموزش قرار می‌گیرد، بنابراین نسبت به روش بهینه سازی تصویر سرعت بیشتری دارد[۳].

«ایان گودفلو و همکارانش» در سال۲۰۱۴ ، با ترکیب دو شبکه مولد و تشخیص‌دهنده، فریم‌ورک شبکه‌ی مخالف مولد را ارائه‌دادند. که در آن شبکه مولد و تشخیص‌دهنده بصورت معکوس آموزش می‌بینند[۱۱]. با موفقیت روش GAN در تولید تصاویر، روش‌های انتقال‌سبک نیز به روش GAN ارائه شده‌اند. روش GAN نیز به روش آفلاین آموزش می‌بینند، بنابراین این روش هم در دسته‌ی روش‌های مبتنی بر بهینه‌سازی مدل قرار می‌گیرد.

انتقال سبک به روش بهینه‌سازی تصویر: «**گیتس و همکارانش**» در سال ۲۰۱۵، اولین شبکه‌ی انتقال سبک عصبی را ارائه دادند. در این روش محتوای عکس بعنوان ویژگی‌هایی از شبکه طبقه‌بندی از قبل آموزش دیده شده مدل‌سازی می‌شود و اثر هنری بعنوان خلاصه‌ای از آمار ویژگی‌ها مدل‌سازی می‌شود. در این روش تصویر با هدف تطبیق توزیع ویژگی‌های CNN مورد نظر، بطور مکرر بهینه‌سازی می‌شود. [۹]، [۳]. «**یانگ چن و همکارانش**»در سال ۲۰۱۷، شبکه‌ای را برای انتقال عکس به کمیک طراحی کردند. [۱۲]. **انتقال سبک به روش Feed-forward:** «**جانسون و همکارانش**» در سال ۲۰۱۶، با الهام از موفقیت «گتیس و همکاران[۹]) »، خطای ادراکی را در شبکه Feed-forward استفاده کردند(قبلا خطای پیکسل استفاده می‌شد) و شبکه را برای انتقال‌سبک و وضوح فوق‌العاده تصویر آموزش دادند. «**شون هوانگ و همکارش**» در سال۲۰۱۷، روشی را برای انتقال‌سبک دلخواه در زمان‌واقعی ارائه دادند. «**ییجون لی و همکارانش**» در سال ۲۰۱۷، روشی را برای انتقال‌سبک دلخواه ارائه دادند. «**دونگ دونگ چن و همکارانش**» در سال ۲۰۱۷، شبکه‌ای را طراحی کردند که انتقال‌سبک عمومی را برای ویدیو انجام می‌دهد. «**دونگ دونگ چن و همکارانش**» در سال ۲۰۱۸، روشی را برای انتقال‌سبک به تصویر یا ویدیوهای استریوسکوپی در زمان‌واقعی ارائه دادند، در واقع روش قبلی‌شان (Chen, Liao, et al.) را برای دو دید چپ و راست گسترش دادند [۱۳]. **انتقال سبک به روGAN:** «**جان یان زو**» در سال ۲۰۱۷، یک فریم‌ورک چرخه‌ای برای شبکه‌های رقابتی برای انتقال‌بافت ارائه دادند که با داده‌های غیر زوجی آموزش می بیند. این شبکه دارای دو مولد و دو تشخیص‌دهنده است[۱۴].

[۱] روش‌هایی است که بافت یک تصویر را تغییر می‌دهند. سنتز بافت،«استخراج و انتقال بافت از تصویر منبع به تصویر هدف است»

«**ترو کراس و همکارانش**» در سال ۲۰۱۸، یک روش جدیدی از GAN را تحت عنوان « GAN افزایشی» معرفی کردند که توانست سرعت و پایداری شبکه را افزایش دهد و تصاویر با کیفیتی با رزولیشن ۱۰۲۴×۱۰۲۴ بسازد. «جری لی» در سال ۲۰۱۸، یک شبکه دوگانه برای ترجمه تصاویر ارائه دادند. «**کویچی هامادا**» در سال ۲۰۱۸، برای تولید کاراکترهای کامل انیمه، یک ساختار شرطی افزایشی را ارائه دادند. «**یانگ چن و همکارانش**» در سال ۲۰۱۹، یک شبکه انتقال‌سبک کارتونی برای چندین سبک با استفاده از داده‌های غیر زوجی ارائه دادند [۱۵]. «**زیجیانگ ژنگ و همکارانش**» در سال ۲۰۱۹، شبکه‌ای را برای تبدیل چهره به کاریکاتور طراحی کردند که بر روی داده‌های غیرزوجی آموزش می‌بینند. «**رویزنگ وو و همکارانش**» در سال ۲۰۱۹، شبکه‌ای را برای تولید چهره کارتونی طراحی کردند. «**سیتو شیانگ و همکاران**» در سال ۲۰۱۹، یک شبکه رقابتی تمایزدهنده محتوا و سبک را برای متمایزسازی سبک‌های مختلف بر اساس نام هنرمند ارائه‌دادند.[16]. «**جی جینگ ژنگ**» در سال ۲۰۱۹، شبکه ای را برای سنتز عکس به طراحی روی داده‌های زوجی آموزش دادند. [۱۷]. «**ونجو زو و همکارانش**» در سال ۲۰۱۹، با افزودن یک خودرمزنگار متغیر به شبکه GAN، یک شبکه برای تولیدات متنوع ارائه دادند. «**یانگ چن و همکارانش**» در سال ۲۰۱۸، یک شبکه‌ی اختصاصی برای ترجمه عکس(صحنه‌های جهان واقعی) به کارتون با داده‌های غیر زوجی با ساختار GAN ارائه دادند. «**یانگ چن و همکارانش**» در سال ۲۰۱۹، شبکه‌ای برای ترجمه ویدیو به کارتون یا انیمیشن در زمان‌واقعی ارائه دادند و توانستند بصورت زنده ترجمه ویدیو به ویدیو را انجام بدهد[۱۸]. «**مکیجپه اسکو و همکارانش**» در سال ۲۰۱۸، شبکه‌ای ارائه‌دادند که فیلم را دریافت می‌کند و خلاصه‌ای از فیلم را به صورت کمیک ارائه می‌دهد[۱۹]. «**شین یو لی و همکارانش**» در سال ۲۰۱۹، یک شبکه اختصاصی، بصورت یک شبکه رقابتی توجهی[1] (AAN) برای کارتونی‌سازی سلفی ارائه دادند که با داده‌های غیر زوجی آموزش داده شده‌است. [۲۰].

۳– پروتوتایپ پیشنهادی

برای ارائه پروتوتایپ بازی بصورت کارتونی و نمایش کارتونی‌سازی فضای بازی، دموی بازی « BMW M2 - Forza Horizon 4[2] » که درمورد شبیه‌ساز رانندگی است، انتخاب شد، علت انتخاب این بازی، انتشار زیاد بازی های تمرین رانندگی است و معمولا اکثر gamer ها حداقل یکبار این نوع بازی را تجربه کرده‌اند. دموی این بازی، بصورت نمایش واقعیت افزوده ارائه شده است،که در کادر بالای سمت چپ دمو، تصویر بازیکن(gamer) نشان داده شده است که gamer بازی را توسط هدایت فرمان و کلاچ و ترمز که بصورت شبیه‌سازی مکانیکی از ماشین واقعی ساخته شداست، انجام می‌دهد. gamer با استفاده از عینک آناگلیف، حضور خود را در فضای بازی احساس می‌کند و کاملا در آن غوطه ور می‌شود و حتی در بعضی موارد، بازیکن لباس‌هایی مانندKor-FX جهت ضربه زدن به بدن و احساس هیجان بیشتر استفاده می‌شود. برای ایجاد این پروتوتایپ، بعد از دانلود دموی بازی، برای نمونه چند فریم آن استخراج شد، سپس این فریم‌ها توسط روشی که «یانگ چن و همکارانش[۱]» ارائه دادند، به تصاویر کارتونی تبدیل یافت. در شکل ۲، نتایج کارتونی‌سازی نمایش داده شده‌است. از نتایج روش«یانگ چن و همکارانش[۱] که بصورت شبکه‌های عصبی مولد رقابتی(GAN) ارائه شده، استفاده شد، به علت اینکه هدف از این مقاله انتشار نتایج کارتونی‌سازی به بازی و مرور روش‌های کارتونی‌سازی است و نتایج روش«یانگ چن و همکارانش[۱] برای چهار سبک کارتونی موجود است.

[1] Attentive Adversarial Network
[2] https://www.youtube.com/watch?v=rtYkboNKThQ

۴– نتایج

در شکل ۳، نتایج کارتونی‌سازی بصورت تصویر نمایش داده شده‌است.

Miyazaki Hayao Style	Hosoda Mamoru Style	دموی بازی BMW M2 - Forza Horizon 4
Shinikai Makoto Style	Paprika Style	

۵– نتیجه گیری

با توجه به مطالبی که در بخش پیشینه ارائه شد، در این مقاله، یک دید کلی از روش‌های انتقال‌سبک ارائه شد. استفاده از فضای کارتونی در بازی از دو جنبه قابل توجه است. ۱– غالبا برای بازی‌های رایانه‌ای از سیستم‌هایی با کارت گرافیک با GPU بسیبار قدرتمند و با حافظه‌ی بالای ۸ گیگا بایت، استفاده می‌شود. فضای کارتونی چون از گرافیک دو بعدی استفاده می‌کند، نیازمند حجم حافظه کمتری است و این حجم حافظه کمتر، باعث در دسترس بودن بازی برای اغلب پلتفرم‌های سخت‌افزاری خواهد شد. ۲– با توجه به نتایجی که در شکل ۳ نشان داده شد و مطالبی که در مورد تاثیر ساده‌سازی عناصر بصری بر انتقال مفاهیم و همچنین انتقال حس بازی هنگام مشاهده‌ی غارنوشته‌ها، در بخش مقدمه توضیح داده شد و همچنین مطالبی که در بخش پیشینه پژوهش، در مورد جام سفالین و وسایل اولیه انیمیشن مطرح شد، قطعا کارتونی سازی در جذابیت بازی تاثیرگذار خواهد بود. علاوه بر این در قسمت پیشینه پژوهش، در مورد روش«دونگ دونگ چن و همکارانش[۱۳]» برای کارتونی‌سازی ویدیو استرسکوپی، توضیح داده شد. این روش، کارتونی‌سازی را برای هر دو دید چپ و راست از طریق عینک آناگلیف انجام داده است. استفاده از عینک آناگلیف، حس حضور در فضای باز را برای gamer فراهم می‌کند و باعث غوطه‌وری بیشتر می‌شود، این حس حضور در فضای واقعی در بازی برای شبیه‌سازهای رانندگی نیز شامل حس و درک سرعت است. نمایش فضای کارتونی توسط عینک آناگلیف، علاوه بر تجربه‌ی جدید حس حضور در فضای کارتونی، حس خیال و انتزاع و توهم را نیز فراهم می‌کند.

۷– پیشنهاد برای تحقیقات آتی

در آینده قصد داریم، تحقیقات بیشتری در مورد کارتونی‌سازی فضای بازی همراه با نظرسنجی از gamer ها ارائه دهیم.

۸– قدردانی

از خانواده‌ام کمال تشکر و قدردانی را دارم. از اساتید محترم مشاور و راهنما برای راهنمایی‌هایشان و از ریاست دانشکده چندرسانه‌ای و همچنین از آقای محمد ببردل بناب نیز کمال تشکر و قدردانی را دارم.

۱۰–مراجع

۱. اصلانی، مژگان، معصومه مرادپور، و زهره نهاردانی. ۱۳۹۱. کارگاه هنر ۲. تهران: موسسه فرهنگی مدرسه برهان(انتشارات مدرسه).
۲. بزرگمهر، مانی. ۱۳۸۹. بزرگترین داستان عالم هنر: پیدایش تصویر متحرک. پرده سینما. (دسترسی در ۱۳۹۹/۰۷/۱۹).
۳. پویانمایی ،۱۳۹۹. https://fa.wikipedia.org/wiki/پویانمایی (دسترسی در ۱۳۹۹/۰۷/۱۹).
٤. توماتروپ. ، ۱۳۹۸.https://fa.wikipedia.org/wiki /توماتروپ (دسترسی در ۱۳۹۹/۰۷/۱۹).

5. *Phenakistiscope. (2020, September 16). Retrieved October 10, 2020, from https://en.wikipedia.org/wiki/Phenakistiscope*
6. *Zoetrope. (2020, October 04). Retrieved October 10, 2020, from https://en.wikipedia.org/wiki/Zoetrope*
7. *Flip book. (2020, October 02). Retrieved October 10, 2020, from https://en.wikipedia.org/wiki/Flip_book Wikipedia. (2020, October 04).*
8. *Animation. (2020, October 05). Retrieved October 09, 2020, from https://en.wikipedia.org/wiki/Animation*

۹. سارانی، محمد. ۱۳۹۱. انیمیشن چیست. *http://bandargraph.blogfa.com/post/16* (دسترسی در ۱۳۹۹/۰۷/۱۹).

10. *Whiteboard animation. (2020, October 11). Retrieved October 14, 2020, from https://en.wikipedia.org/wiki/Whiteboard_animation*
11. *Animation Styles & Types. (2020, July 28). Retrieved January 06, 2021, from https://biteable.com/blog/animation-styles-types/*
12. *Goodfellow, I., Bengio, Y., & Courville, A. (2017). Deep learning. Cambridge, MA: The MIT Press.*

۱. Chen, Y., Y.-K. Lai, and Y.-J. Liu. *Cartoongan: Generative adversarial networks for photo cartoonization*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018.
۲. Wibisana, F., H. Reynaldi, and S. Sn, *STUDI KASUS DAN ANALISA METODE TEKNIK ROTOSCOPY PADA FILM. Finlandia Wibisana: 156020061*. 2019, Fotografi & Film.
۳. Jing, Y., et al., *Neural style transfer: A review*. IEEE transactions on visualization and computer graphics, 2019.
٤. *Deep Learning*. [cited 2020 Retrieved June 20, 2020]; Available from: https://www.deeplearningbook.org./
٥. Kyprianidis, J.E., et al., *State of the" Art": A Taxonomy of Artistic Stylization Techniques for Images and Video*. IEEE transactions on visualization and computer graphics, 2012. **19** :(٥)p. 866-885.
٦. Buck, I., et al., *Performance-driven hand-drawn animation*, in *ACM SIGGRAPH 2006 Courses*. 2006. p. 25-es.

.۷ Kyprianidis, J.E. and J. Döllner. *Image abstraction by structure adaptive filtering*. in *TPCG*. 2008.

.۸ Luque, R.R., *The cel shading technique*. 2012, Citeseer.

.۹ Gatys, L.A., A.S. Ecker, and M. Bethge, *A neural algorithm of artistic style.* arXiv preprint arXiv:1508.06576, 2015.

.۱۰ Johnson, J., A. Alahi, and L. Fei-Fei. *Perceptual losses for real-time style transfer and super-resolution*. in *European conference on computer vision*. 2016. Springer.

.۱۱ Goodfellow, I., et al. *Generative adversarial nets*. in *Advances in neural information processing systems*. 2014.

.۱۲ Chen, Y., Y.-K. Lai, and Y.-J. Liu. *Transforming photos to comics using convolutional neural networks*. in *2017 IEEE International Conference on Image Processing (ICIP)*. 2017. IEEE.

.۱۳ Chen, D., et al. *Stereoscopic neural style transfer*. in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018.

.۱٤ Zhu, J.-Y., et al. *Unpaired image-to-image translation using cycle-consistent adversarial networks*. in *Proceedings of the IEEE international conference on computer vision*. 2017.

.۱٥ Chen, Y., et al. *CartoonRenderer: An Instance-based Multi-Style Cartoon Image Translator*. in *International Conference on Multimedia Modeling*. 2020. Springer.

.۱٦ Xiang, S. and H. Li, *Disentangling style and content in Anime illustrations.* arXiv preprint arXiv:1905.10742, 2019.

.۱۷ Zheng, J., et al., *Feature Encoder Guided Generative Adversarial Network for Face Photo-Sketch Synthesis.* IEEE Access, 2019. **7**: p. 154971-154985.

.۱۸ Chen, Y., et al. *Animating Your Life: Real-Time Video-to-Animation Translation*. in *Proceedings of the 27th ACM International Conference on Multimedia*. 2019.

.۱۹ Pęśko, M., et al., *Comixify: Transform Video Into Comics.* Fundamenta Informaticae, 2019. **168**(2-4): p. 311-333.

.۲۰ Li, X., et al. *Everyone is a Cartoonist: Selfie Cartoonization with Attentive Adversarial Networks*. in *2019 IEEE International Conference on Multimedia and Expo (ICME)*. 2019. IEEE.