

## رفع اعوجاج و بازسازی کد های QR ناخوانا با استفاده از شبکه‌ی

### عصبی عمیق

میلاذ منفرد<sup>۱</sup> و عباس کوچاری<sup>۲</sup>

#### چکیده

امروزه بارکد ها نقش پررنگی در صنایع مختلف ایفا می کنند و در بین بارکد های دوبعدی معروف ترین آن ها یعنی کد QR (کد پاسخ سریع) رشد روز افزونی داشته است. هدف اصلی این مقاله ارائه ی روش رفع نویز مبتنی بر شبکه ی عصبی عمیق خود شناس است که با استفاده از آن بتوان QR های مخدوش غیر خوانا را دوباره به حالت خوانا بازگرداند. برای ایجاد نویز و اعوجاج بر خلاف مقالات دیگر که از شبیه نویز استفاده و آن را به تصویر اضافه کردند، از چالش استخراج QR نهان نگاری شده درون یک تصویر رنگی استفاده شده تا با جمع آوری داده هایی از جهان واقعی، ارزیابی جامع تر و واقعی تری از کیفیت و خوانایی QR های رفع نویز شده با روش ارائه شده را داشته باشیم. در نتیجه دیتا ست جامعی از QR های مخدوش حاصل از سه رویکرد استخراج نهان نگاره متفاوت بعد از حمله ی اسکرین-دوربین را ایجاد نمودیم. برای فرایند رفع نویز نیز سه شبکه ی MCNN مستقل برای هر یک از سه رویکرد استخراج استفاده شده است که ارتقاع یافته از شبکه ی U-net می باشد.

#### کلیدواژه‌ها

حذف نویز، بازسازی کد QR، شبکه ی عصبی خود شناس، MCNN، یادگیری عمیق، یادگیری ماشین

#### ۱ مقدمه

حوزه هایی مانند صنعت حمل نقل، بیزینس، تولید، لوازم خانگی و غیره استفاده شده است. امروزه فناوری بارکد تک بعدی به بلوغ خود رسیده و سهم زیادی در افزایش سرعت جمع آوری داده و پردازش اطلاعات دارد. ازسوی دیگر بارکد های تک بعدی سنتی دارای محدودیت هایی مانند حجم اندک برای ذخیره سازی داده، عدم توانایی برای تصحیح خطا و وابستگی به پایگاه داده می-باشند. [۱]

فناوری بارکد نقش پررنگی در تکنولوژی های جمع آوری داده و شناسایی خودکار ایفا می کند. این فناوری ها شامل بخش هایی مانند کلگذاری، نمایش نمادین، چاپ، شناسایی، جمع آوری و پردازش داده ها می باشند. بارکد یک بعدی به صورت گسترده در

برای جبران کاستی های فوق، بار کد های دوبعدی در اوایل دهه ی ۹۰ میلادی پدیدار شدند. نمونه ی بسیار معروف و همه گیر تر از آن ها QR نام دارند که شکل ۱ نمونه هایی از کاربرد های بسیار گسترده ی آن در ذخیره سازی اطلاعاتی مانند مختصات و نقشه، آدرس وب، دستور پرینت، مشخصات و قیمت محصول و مشخصات فردی و آدرس پست الکترونیکی مشاهده می شود.

این مقاله در تیرماه ۱۴۰۰ دریافت، در آبان ماه بازنگری و در آذرماه همان سال پذیرفته شد.

<sup>۱</sup> دانشجوی کارشناسی ارشد گرایش هوش مصنوعی، دانشگاه آزاد واحد علوم و تحقیقات، دانشکده ی مکانیک-برق-کامپیوتر

رایانامه: [miladmonfared74@gmail.com](mailto:miladmonfared74@gmail.com)

<sup>۲</sup> دانشگاه آزاد واحد علوم و تحقیقات، دانشکده ی مکانیک-برق-کامپیوتر

رایانامه: [koochari@srbiau.ac.ir](mailto:koochari@srbiau.ac.ir)

نویسنده مسئول: میلاذ منفرد

DOR: 20.1001.1.23831197.1401.9.3.6.1

گیری‌ها را پیش‌بینی یا شبیه‌سازی کنیم. اما بازیابی پارامترهای توصیف‌کننده‌ی وضعیت یک سیستم از اندازه‌گیری‌های به‌عمل آمده نیازمند حل یک مسئله‌ی معکوس است، که غالباً به‌صورت قطعی یا غیرتکراری قابل حل نمی‌باشند.

مسائل معکوس از مهمترین مسائل ریاضی در علوم هستند. با این حال، در بیشتر موارد، راه‌حل‌های ارائه‌شده آنها برای حل مشکلاتی تحت شرایط خاص و استفاده مناسب از دانش خاص طراحی شده‌اند و نمی‌توان راه‌حل اینگونه مسائل را به سایر مسائل معکوس تعمیم داد. ماهیت غیرخطی بسیاری از این مسائل تاکنون از شکل‌گیری چارچوبی به‌عنوان یک حل‌کننده عمومی جلوگیری کرده است. مسئله‌ی رفع نویز به‌صورت نگاشت تصویر به تصویر یکی از نمونه‌های چالش‌برانگیز اینگونه مسائل می‌باشد. راهکارهای متعددی برای حل مسائل رفع نویز بر روی تصاویر ارائه شده‌اند اما با انگیزه پیشرفت‌های اخیر در توسعه سخت‌افزار محاسبات و هدایت شده توسط مجموعه‌های داده بزرگ [۲، ۳]، یادگیری عمیق [۴] اخیراً پتانسیل بالایی را به ویژه در کارهایی مانند نگاشت تصویر به تصویر، سوپر رزولوشن [۵]، رفع نویز تصویر [۶] و تولید تصویر [۷] و موارد دیگر نشان داده است.

در بسیاری از مسائل معکوس [۸، ۹] استفاده از یادگیری عمیق به نتایج امیدوارکننده‌ای منتهی شده است، قضیه تقریب جهانی<sup>۲</sup> تضمین می‌کند که شبکه‌های عصبی می‌توانند عملکردهای دلخواه را به خوبی تخمین بزنند [۱۰-۱۲]. با این حال، هیچ تضمینی برای پیدا کردن وزن‌های شبکه عصبی با بهینه‌سازی وجود ندارد و به دلیل غیرخطی بودن بالای چنین معماری‌های عمیقی، عملکرد آنها به شدت به تنظیمات مناسب‌های پارامترها بستگی دارد. حتی مدل‌های بزرگ و پیشرفته اخیر نیز نیاز به سالها توسعه و ده‌ها هزار دلار هزینه برای آموزش دارند [۱۳، ۱۴]. به همین دلیل اهمیت آموزش اینگونه شبکه‌ها با داده‌های واقعی و کاربردی حائز اهمیت است. همچنین ترندهای زیادی برای تنظیم پارامترهای یک شبکه عصبی مانند تنظیم مناسب نرخ یادگیری، ریگولایزاسیون، طراحی لایه‌های پنهان، کرنل‌های کانولوشن و معماری شبکه وجود دارد [۱۵، ۱۶]، با این وجود هیچ‌یک از این اقدامات نمی‌تواند هم‌زمان هم سرعت همگرایی و هم کیفیت محصول حاصل را افزایش دهد [۱۷].

در یادگیری ماشین، شبکه‌ی خودشناس<sup>۳</sup> یک شبکه‌ی عصبی پیشروی<sup>۴</sup> سه لایه است که برای بازسازی ورودی خود آموزش دیده

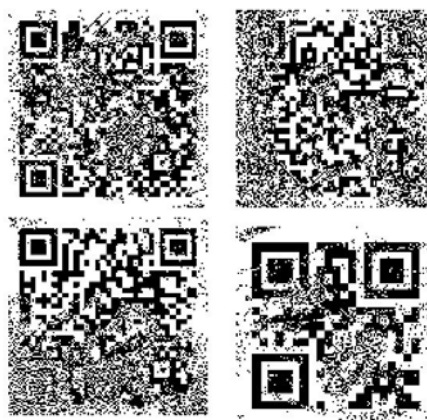
مشخص می‌شود که این بارکد‌ها ذخیره‌سازی داده، تصحیح خطا و عدم وابستگی به پایگاه‌داده را به خوبی پوشش می‌دهند. در بین بارکد‌های دوبعدی، کد QR توانایی بهتری در پردازش کدگذاری‌ها در کاراکترهای غیرانگلیسی و اطلاعات تصویری دارد. علاوه بر این کد QR نسبت به دیگر بارکد‌های دوبعدی از خوانش بسیار سریع و تشخیص همه‌جانبه بهره‌مندی برد.



شکل ۱ حوزه‌های کاربردی کدهای QR

در بسیاری از موارد ممکن است کدهای QR دچار اعوجاج و نویز شده و خوانایی خود را از دست بدهند. برای مثال بعد از درون‌نگاری و استخراج یک QR درون تصویر رنگی، شانس ناخوانا بودن آن به علت اعوجاج‌های هندسی و عدم استخراج صحیح تمامی پیکسل‌های QR بالا است. در شکل ۲ نمونه‌ای از QRهای ناخوانا بعد از فرایند استخراج از تصویر را مشاهده می‌کنید.

هدف از انجام این تحقیق، ارائه‌ی روشی قابل اطمینان برای رفع اعوجاج و نویزهای QRهای ناخوانا و تبدیل دوباره آنها به QRهای خوانا می‌باشد.



شکل ۲ QRهای نویزی و مخدوش بعد از استخراج از تصاویر نهان‌نگاری شده با رویکرد ساده

1 translation

2 Universal approximation theorem

3 Autoencoder Neural Network

4 Feed-Forward

بیشتر مسائل فیزیکی با توجه به توصیف کامل وضعیت یک سیستم به ما امکان پیش‌بینی می‌دهند. ما می‌توانیم برخی از اندازه

در ادامه مقاله، در بخش دوم شرح مختصری از سیستم نهان نگاری و استخراج QR به منظور جمع آوری داده های نویزی داده می شود، در بخش سوم به شرح مجموعه داده و چگونگی جمع آوری آن پرداخته می شود، در بخش چهارم روش ارائه شده بر مبنای شبکه ی عصبی خودشناس برای رفع اعوجاج های QR توضیح داده خواهد شد. نتایج تجربی در بخش ششم ارائه شده و نهایتاً در بخش آخر نتیجه گیری ارائه می شود.

## ۲ استخراج QR های مخدوش از تصویر نهان نگاری شده

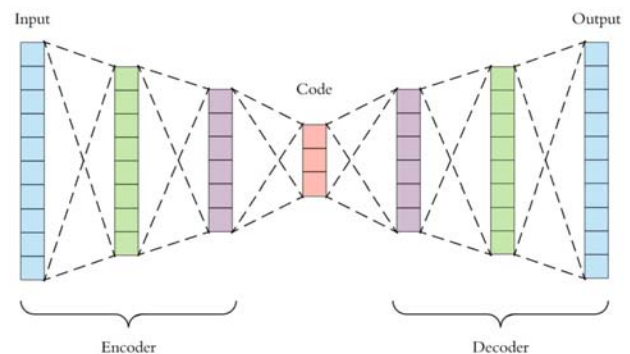
متاسفانه تحقیقات انجام گرفته در زمینه ی رفع اعوجاج های کد های QR به دلیل کمبود داده های واقعی مورد توجه کمی قرار گرفته اند. از این رو در این مقاله برای استفاده ی کاربردی از روش ارائه شده در رفع نویز QR های مخدوش، از یکی از پژوهش های پیشین در حوزه ی نهان نگاری QR مقاوم به فرایند اسکرین-دوربین در تصاویر رنگی استفاده شده است.

در پژوهش های معمول در حوزه ی نهان نگاری تصویر معمولاً یک لوگو درون تصویر میزبان درون نگاری می شود و بعد از استخراج با وجود نویز های ایجاد شده بر روی لوگوی استخراج شده، باز هم لوگو قابل تشخیص است، اما استخراج کد های QR بعد از حملات سنگینی مانند حمله ی اسکرین-دوربین با چالش های فراوانی رو به روست. حمله ی اسکرین-دوربین که جزو حملات جدید در این حوزه می باشد، خود شامل حملات گوناگونی است که علاوه بر ایجاد نویز، در بسیاری از موارد محتوای درون QR را نیز دچار تغییر می کند و عملاً کد QR قابلیت اسکن خود را از دست می دهد. فاکتور های بسیاری در ایجاد شدن این نوع اعوجاج ها موثر اند مانند کیفیت تصویر گرفته شده، کیفیت نمایشگر، فاصله دوربین، نوع نگاشت دوربین، زاویه لنز و غیره، اما هدف نهایی قابلیت اسکن و خوانایی QR های استخراج شده از تصاویر درون نگاری شده است.

در پژوهش اشاره شده، روش نهان نگاری مبتنی بر تبدیل فوریه چهارگانه [۱۸، ۱۹] بر روی تصاویر رنگی می باشد و درون نگاری و استخراج QR ها بر مبنای سه رویکرد ساده، تکرار چهارگانه و رای گیری شده طراحی شده است. لازم به ذکر است که فرایند استخراج نهان نگاره های QR بعد از عکس گرفتن از روی تصویر نهان نگاری شده که بر روی نمایشگر در حال پخش است انجام گرفته و تمامی مراحل استخراج، بر روی تصویری که با دوربین تلفن همراه گرفته شده است انجام می گیرد. تصویر برداری از روی نمایشگر حاوی تصویر نهان نگاری شده و اقدام برای استخراج QR درون نگاری شده در آن، خود باعث بسیاری

است به طور معمول، یک لایه ی ورودی، یک لایه ی پنهان (به اصطلاح تنگنا یا فضای پنهان) و یک لایه ی خروجی دارد. خودشناس ها یک الگوریتم یادگیری بدون ناظر هستند؛ بنابراین برچسب ها یا در دسترس نیستند یا ما به آنها اهمیت نمی دهیم. خودشناس ها به جای آموزش شبکه جهت پیشبینی خروجی از مقدار ورودی، برای بازسازی ورودی خود آموزش می بینند و تلاش می کنند تا ورودی خود را با حداقل انحراف به خروجی کپی نمایند و بدین منظور واضح است که ابعاد ورودی و خروجی باید یکسان باشند. بنابراین، خودشناس ها دو نگاشت را یاد می گیرند: اول از ورودی به لایه پنهان (این نگاشت شناسه گذار<sup>۱</sup> نامیده میشود) و دوم از لایه پنهان به خروجی (این نگاشت یک شناسه گشا<sup>۲</sup> نامیده میشود).

روش کار به این صورت است که شناسه گذار، داده ها را از فضای ورودی به فضای مخفی نگاشت می کند و شناسه گشا داده های رمز شده را به فضای اصلی نگاشت می کند. میتوان بازنمایی لایه پنهان را به عنوان یک کد برای بازنمایی ورودی در نظر گرفت، که باید بازنمایی های خوبی از ورودی را بیاموزد و از آن می توان داده های اصلی را به صورت شناسه گشایی باز کرد. البته، کد باید ویژگی های اصلی داده را ضبط کند، در غیر این صورت استخراج چیزهای معنی دار از آن غیر ممکن خواهد بود. در شکل ۳ ساختار یک خودشناس با چند لایه پنهان، نمایش داده شده است.



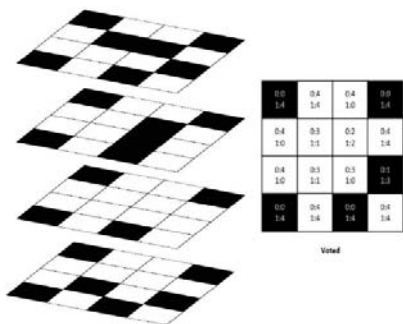
شکل ۳ ساختار یک شبکه ی عصبی خودشناس

اگر هدف خودشناس ها تنها کپی ورودی در خروجی بود، آنها غیر قابل استفاده بودند؛ اما با آموزش یک خودشناس برای کپی ورودی در خروجی، لایه پنهان بازنمایی ویژگیهای مفیدی را در بر خواهد گرفت. همچنین برای آموزش شبکه ی خودشناس رفع نویز باید به جای قرار دادن تصاویر همسان به عنوان داده های آموزشی، از تصاویر بدون نویز را به عنوان هدف آموزشی و تصاویر نویزی متناظر به عنوان ورودی استفاده شود.

<sup>1</sup> Bottleneck

<sup>2</sup> Encoder

<sup>3</sup> Decoder

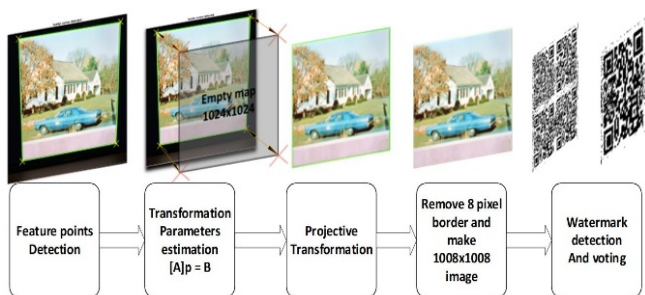


شکل ۵ شمای کلی رای گیری پیکسلی

همانطور که در شکل ۴ قابل مشاهده است، این تکنیک از نویزها سطحی می‌کاهد اما هنوز QR قابل خوانش نمی‌باشد و دچار اعوجاج‌های عمیقی است.

## ۲-۲ استخراج

در این بخش به صورت کلی روش و مراحل استخراج تصویر QR نویزی بیان شده است. شکل ۵ شمای کلی استخراج نهان نگاره‌ی QR را نشان می‌دهد.



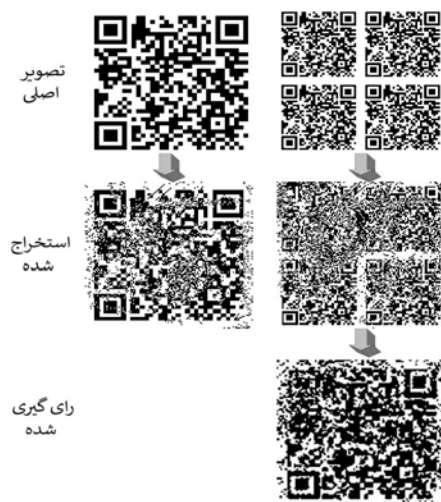
شکل ۶ شمای کلی استخراج نهان نگاره از تصویر گرفته شده با دوربین از روی نمایشگر

به صورت خلاصه در فاز استخراج نهان نگاره از روی تصویر گرفته شده با دوربین، ابتدا چهار نقطه‌ی گوشه‌ی تصویر را با استفاده از حاشیه‌ای که آن را احاطه کرده است شناسایی می‌کنیم، سپس با استفاده از عکس، عملیات تبدیل پروجکتیو<sup>۱</sup> و یافتن پارامترهای تبدیل، نگاشت تصویر به تصویر اصلاح شده و بدون اعوجاج‌های هندسی ناشی از فرایند اسکرین-دوربین صورت می‌گیرد. در گام بعدی حاشیه‌ی اطراف تصویر اصلاح شده را حذف می‌کنیم. حال که تصویر پیش‌رو کاملاً بدون اعوجاج‌های هندسی و از نظر پیکسل به پیکسل با تصویر اصلی همگام است، با انجام عملیات مبتنی بر بلاک و تبدیل فوریه‌ی چهارگانه بر روی هر بلاک تصویر نهان نگاره‌ی شده‌ی QR را استخراج می‌نماییم. در گام نهایی در صورت استخراج رویکرد درون نگاره‌ی چهارگانه می‌توانیم از چهار QR به صورت پیکسل به پیکسل رای گیری کرده

از اعوجاج‌ها و نویزهایی مانند اعوجاج‌های هندسی می‌شود که یکی از اهداف اصلی این مقاله، مقابله با آن‌هاست. در بخش ابتدایی این قسمت به انواع رویکرد‌های استخراج پرداخته و در بخش بعدی نگاهی گذرا به مرحله‌ی استخراج خواهیم داشت.

## ۱-۲ رویکرد‌های استخراج

همان‌طور که پیش‌تر اشاره شد، سه رویکرد استخراج QR با نام‌های استخراج ساده، استخراج چهارگانه و استخراج رای گیری شده برای آزمایش و ارزیابی و مقایسه مورد استفاده قرار گرفتند. رویکرد استخراج ساده حاصل درون نگاره‌ی و استخراج یک تصویر QR منفرد درون یک تصویر رنگی می‌باشد اما در رویکرد تکرار چهارگانه به جای درون نگاره‌ی یک QR منفرد، از تکرار ۲ در ۲ آن در یک تصویر استفاده شده. استخراج رای گیری شده نیز حاصل رای گیری بین چهار تصویر QR درون استخراج چهارگانه می‌باشد، به صورتی که بعد از استخراج نهان نگاره اگر از رویکرد تکرار چهارگانه استفاده شده باشد، با استفاده از تکنیک رای گیری پیکسل‌ها، یک تصویر واحد با اعوجاج نسبتاً کم‌تر ایجاد می‌شود که آن را تصویر رای گیری شده می‌نامیم. در شکل ۴ نمونه‌ای از رویکرد ساده، تکرار چهارگانه و رای گیری شده را مشاهده می‌کنید.



شکل ۴ نمونه‌ای از رویکرد درون نگاره ساده و تکرار چهارگانه و استخراج رای گیری بعد از حمله‌ی اسکرین-دوربین

برای ایجاد تصویر رای گیری شده از رویکرد تکرار چهارگانه، بعد از دودویی کردن تصویر QR چهارگانه، چهار تصویر تکرار شونده را با سایزهای یکسان برش داده، سپس تمامی پیکسل‌های هر چهار تصویر را در نظر گرفته و ۰ یا ۱ بودن آن‌ها را چک می‌کنیم. اگر پیکسل در مکان مربوطه در بیشتر از دو تصویر ۱ بود، در تصویر نهایی ۱ قرار می‌دهیم در غیر این صورت ۰. در شکل ۵ نمونه‌ی کوچکی از رای گیری را مشاهده می‌نمایید.

<sup>1</sup> Inverse projective transform



شكل ۷ تمامی تصاویر میزبان استفاده شده و چهار نمونه از انواع سطوح کد های کیو آر مورد استفاده شده است

تا تصویر یکتای حاصل از رای گیری را هم در آزمایشات خود لحاظ کنیم. با در نظر گرفتن سه رویکرد استخراج ، نیاز به سه مجموعه داده می باشد تا به صورت مجزا سه شبکه ی خودشناس را آموزش داده و نتایج را مقایسه و ارزیابی کنیم. در قسمت بعدی به چگونگی جمع آوری مجموعه داده و مشخصات آن ها می پردازیم.

### ۳ مجموعه داده

از آنجایی که تعداد مقالات رفع نویز و اعوجاج های عمیق کد QR اندک بوده و دیتاستی برای QR های نویزی حاصل از استخراج تحت حملات ناشی از فرایند اسکرین-دوربین یافت نمی شد ، مجبور به ایجاد سه مجموعه داده ی مستقل برای QR های نویزی شدیم [۲۰] تا سه شبکه ی عصبی آموزش دیده ی مستقل برای هر سه رویکرد درون نگاری ساده ، تکرار چهارگانه و استخراج رای گیری شده داشته باشیم.

در این مقاله برای ارزیابی و ایجاد داده های آموزش و تست مورد نیاز شبکه ی عصبی عمیق ، از ۳۰ تصویر میزبان برای درون نگاری، ۵۰ کد QR برای فاز آموزش و ۲۵ کد QR برای فاز تست استفاده شده است که بعد از درون نگاری ، عکس گرفتن با دوربین های مختلف از روی نمایشگر های مختلف و در شرایط نوری مختلف و سپس استخراج QR ها ، مجموعه داده ای با ۱۵۰۰ داده ی آموزشی QR نویزی و ۷۵۰ داده ی تست QR نویزی با رزولوشن ۵۱۲ در ۵۱۲ پیکسل گرد آوری شد . با در نظر گرفتن سه رویکرد درون نگاری ساده و تکرار چهارگانه و استخراج رای گیری شده در مجموع سه دیتا ست ۲۲۵۰ تایی از انواع QR های نویزی ایجاد کردیم. در شکل ۷ تمامی تصاویر میزبان استفاده شده و چهار نمونه از انواع سطح QR از نظر تراکم و حجم داده را مشاهده می کنید.

برای تست خوانایی QR های استخراج شده یا رفع نویز شده از کتابخانه ی pyzbar تحت زبان پایتون استفاده شده است. تصاویر درون نگاری شده با رویکرد های مختلف را به صورت جداگانه استخراج کرده و با کتابخانه ی معرفی شده سعی به خواندن QR استخراج شده و رفع نویز شده می کنیم. در انتها همانطور که در فرمول ۱ مشخص است، بر اساس تعداد QR های خوانده شده کل نمونه ها، درصد خوانایی را ارزیابی می کنیم.

$$QR\ readability = (read\ samples/total\ samples)*100 \quad (1)$$

### ۴ رفع نویز با استفاده از شبکه ی عصبی عمیق

مسائل معکوس در پردازش تصویر ، تصویربرداری فاز و بینایی ماشین از نظر ساختار نگاشت تصویر ورودی به تصویر خروجی غالباً همانند هستند. اما معمولاً توسط الگوریتم های مختلفی حل می شوند که در این میان شبکه های عصبی عمیق کانولوشن پتانسیل بالایی را برای بسیاری از کاربردها در حوزه ی تصویر پردازی نشان داده اند. اما گاهی اوقات ممکن است آموزش آنها به دلیل غیر خطی بودن ذاتی آن ها چالش برانگیز باشد.

یک شبکه ی عصبی خودشناس نوعی ساختار از شبکه ی عصبی می باشد که از مفهوم انتشار<sup>۱</sup> و پس انتشار<sup>۲</sup> استفاده می نماید. شبکه ی عصبی خودشناس سیستمی است که از نگاشت قطعی برای نگاشت ورودی x به لایه ی پنهان h استفاده می کند.

$$Encoder : y = s(Wx + b) \quad (2)$$

که در رابطه ی ۲، s یک تابع غیر خطی مانند سیگموئید<sup>۳</sup> و رلو<sup>۴</sup> می باشد. مقادیر کد شده ی y با رابطه ی ۳ به مقادیر نوسازی شده ی z دوباره نگاشت می شوند.

$$Decoder : z = s(W'y + b') \quad (3)$$

پارامتر های w و w' ، b و b' با رویکرد کمینه سازی تابع هزینه بهینه می شوند. خودشناس حذف نویز (DAE<sup>۵</sup>) شبکه ی تعمیم یافته ای از خودشناس های کلاسیک می باشد که در آن ما مدل را به یادگیری بازسازی ورودی های مخدوش می کنیم [۲۱]. آموزش خودشناس های رفع کننده ی نویز همانند آموزش یک خودشناس استاندارد از مفاهیم انتشار و پس انتشار استفاده می کند ، البته با لحاظ کردن تفاوت در خروجی ها. در خودشناس های استاندارد خروجی شبکه نگاشتی برابر با ورودی است بنا بر این تفاوتی میان داده های ورودی و داده های لیبل آموزش وجود

<sup>1</sup> Propagation

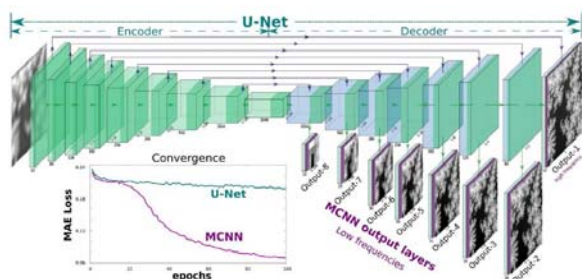
<sup>2</sup> Backpropagation

<sup>3</sup> Sigmoid

<sup>4</sup> ReLu

<sup>5</sup> Denoising Auto Encoder

[۳۰]. این معماری سرعت همگرایی بالاتری برای شبکه را فراهم آورده در عین حال در نقاط بهینه‌ی محلی نمی‌ماند و به صورت کلی روند آموزش را به طور چشمگیری تثبیت می‌کند. همانطور که در پایین سمت چپ شکل ۸ نشان داده شده است. با تنظیمات یکسان، هنگام حل یک مسئله معکوس لاپلاسیان از یک تقریب گرادیان مرتبه دوم، MCNN به سرعت به میانگین خطای مطلق ( $MAE^5$ ) در حدود  $0.07$  می‌رسد در حالی که U-Net معمولی برای آموزش چالش برانگیز است، و در حدود  $0.2$  گرفتار می‌شود. لایه‌های نرمالیزیشن اضافی نیز می‌توانند همگرایی [۳۱]، [۳۲] را تسریع کنند، اما به احتمال زیاد نویزهای نامطلوب را ایجاد می‌کنند و باید با دقت استفاده شوند.



شکل ۸ با همگام شدن خروجی‌های تمامی فرکانس‌ها، شبکه‌ی MCNN عملکرد بهتری نسبت به شبکه‌ی U-net نشان می‌دهد. در تصویر فوق ساختار یک شبکه‌ی MCNN نشان داده شده است. شبکه‌ی MCNN از یک شبکه‌ی U-Net (قسمت بالا) و اضافه شدن ۷ لایه‌ی مولتی‌رزولوشن (از خروجی ۲ تا خروجی ۸) تشکیل شده است. با این روش تصویر ورودی ابتدا به تنسورهای بعد بالاتر تبدیل شده و سپس در فاز شناسه‌گشا به ۸ تصویر با رزولوشن‌های مختلف دیکد میشوند تا فرکانس‌های مختلف را سریع‌تر ایجاد کند. همانطور که در منحنی همگرایی داده‌های تست نشان داده شده است، شبکه‌ی MCNN برتری چشمگیری نسبت به U-Net کلاسیک دارد. [۲۶]

در جدول ۱ معماری لایه‌های شبکه‌ی خودشناس MCNN را مشاهده می‌کنید. معماری فوق بر اساس اضافه کردن لایه‌های مخفی چند وضوح به معماری پایه‌ی شبکه‌ی U-net ایجاد شده. در این پژوهش به صورت کلی از ساختمان شبکه‌ی MCNN استفاده شده است اما آموزش شبکه از پایه با استفاده از دیتاست جمع‌آوری شده انجام گرفته است. در برخی موارد جزئی‌تر مانند کرنل ریگولاریزیشن L2 و پارامتر  $\lambda$  و همچنین در سایز batch normalization تغییراتی اعمال شده است. در تمامی لایه‌های کانولوشنال از کرنل‌های تنظیم‌کننده<sup>۶</sup> L2 با پارامتر  $\lambda$  برابر  $1.0e-3$  استفاده شده است.

ندارد. اما در خودشناس‌های رفع نویز ورودی‌های نویزی در برابر لیل‌های غیر نویزی خودشان قرار خواهند گرفت تا شبکه مجبور به یادگیری نوع نویز و بازسازی تصویر سالم شود. در بسیاری از پژوهش‌ها [۲۲] برای ایجاد داده‌های آموزشی، نویزهای شبیه‌سازی شده مانند نویز گوسی را به تصاویر سالم اضافه کرده و مجموعه داده‌ی ورودی خود را می‌سازند. ما در این پژوهش برای ایجاد داده‌های نویزی از روش ارائه شده مبتنی بر استخراج نهان نگاره‌ی QR استفاده کرده و داده‌های آموزشی خود را آزمایش می‌کنیم.

با آگاهی از این واقعیت که یک شبکه عصبی کانولوشن عمیق (DCNN)<sup>۱</sup> اغلب به سرعت اجزای با فرکانس پایین و پس از آن اجزای با فرکانس بالا را به روشی نسبتاً کند [۲۳، ۲۴] بازیابی می‌کند، و همچنین با الهام از ایده‌های روشهای multi-grid [۲۵]، wang و همکارانش [۲۶] شبکه عصبی کانولوشن عمیق چند وضوح‌ی جدید (MCNN)<sup>۲</sup> را پیشنهاد می‌دهند.

در تحقیق به عمل آمده توسط آن‌ها، معماری جدیدی از شبکه‌ی عصبی عمیق کانولوشنال ارائه شده است که سریع‌تر از مدل U-net که مدل پایه‌ی تحقیق آن‌ها بوده است همگرا میشود. از این شبکه‌ی عصبی برای اهدافی مانند شفاف کردن تصاویر میکروسکوپی و بازیابی جبهه‌های موج از اندازه‌گیری مستقیم شدت<sup>۳</sup> استفاده شده است. ما نیز تصمیم گرفتیم که از ساختار این شبکه استفاده کرده و با مجموعه داده‌ی از تصاویر QR های سالم و نویزی که خود به جمع‌آوری آن اقدام کردیم، از ابتدا به آموزش آن پرداخته و نتایج را تحلیل کنیم، سپس نتایج بدست آمده از آموزش ساختار شبکه‌ی عصبی خودشناس MCNN را با خروجی‌های آموزش ساختارهای استانداردی مانند شبکه‌ی خودشناس<sup>۴</sup> CDAE [۲۲] و U-net بر روی مجموعه داده‌ی خود مقایسه کنیم.

همانطور که در شکل ۸ قابل مشاهده است، در معماری شبکه‌ی MCNN با اتصال لایه‌های پنهان کانولوشنال اضافی به ساختار شبکه‌ی U-net [۲۷، ۲۸] در بخش شناسه‌گشا، این امکان فراهم شده است تا در خروجی به تولید هرم رزولوشنی از تصویر مورد نظر پرداخته و به بیان دیگر بتوان خروجی‌هایی از جزئیات درشت‌تر تا ریزتر را ایجاد کرد. اصلاحات بیشتری مانند اضافه کردن لایه‌های مشابه در قسمت شناسه‌گذار نیز صورت گرفت اما تاثیر قابل توجهی در همگرایی شبکه مشاهده نشد. یکی دیگر از تغییرات، تغییر شکل تدریجی معماری با شروع آموزش اجزای فرکانس‌های پایین در مراحل اولیه، سپس قرار دادن لایه‌های جدید برای تطابق با ویژگی‌های فرکانس بالا در طول آموزش است، که در برخی از کاربردهای اخیر نشان داده شده است [۲۹]،

<sup>1</sup> Deep Convolutional Neural Network

<sup>2</sup> Multi-resolution Convolutional Neural Network

<sup>3</sup> recovering wavefronts from direct intensity measurements

<sup>4</sup> Convolutional Denoiser Auto Encoder

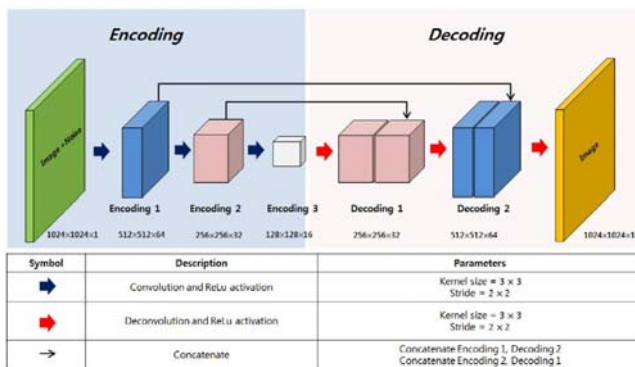
<sup>5</sup> Mean Absolute Error

<sup>6</sup> Regularizes

جدول ۱ ساختار لایه های شبکه ی خودشناس MCNN

L15->L16: Deconvolutional and ReLu activation	Filter size: $3 \times 3 \times 192$ Stride: $2 \times 2$ Padding: same
L18: Concatenate L16 , L3	-
L18->Last64: Convolutional and ReLu activation (MCNN Layer)	Filter size: $3 \times 3 \times 8$ Stride: $1 \times 1$ Padding: same
Last64->Output-4: Convolutional and Sigmoid activation (MCNN Layer)	Filter size: $3 \times 3 \times 1$ Stride: $1 \times 1$ Padding: same
L18->L19: Deconvolutional and ReLu activation	Filter size: $3 \times 3 \times 128$ Stride: $2 \times 2$ Padding: same
L21: Concatenate L19 , L2	-
L21->Last128: Convolutional and ReLu activation (MCNN Layer)	Filter size: $3 \times 3 \times 8$ Stride: $1 \times 1$ Padding: same
Last128->Output-3: Convolutional and Sigmoid activation (MCNN Layer)	Filter size: $3 \times 3 \times 1$ Stride: $1 \times 1$ Padding: same
L21->L22: Deconvolutional and ReLu activation	Filter size: $3 \times 3 \times 64$ Stride: $2 \times 2$ Padding: same
L24: Concatenate L22 , L1	-
L24->Last256: Convolutional and ReLu activation (MCNN Layer)	Filter size: $3 \times 3 \times 8$ Stride: $1 \times 1$ Padding: same
Last256->Output-2: Convolutional and Sigmoid activation (MCNN Layer)	Filter size: $3 \times 3 \times 1$ Stride: $1 \times 1$ Padding: same
L24->L25: Deconvolutional and ReLu activation	Filter size: $3 \times 3 \times 32$ Stride: $2 \times 2$ Padding: same
L27: Concatenate L25 , init	-
L27->LL: Convolutional and ReLu activation	Filter size: $3 \times 3 \times 16$ Stride: $1 \times 1$ Padding: same
LL->LL : Convolutional and ReLu activation	Filter size: $3 \times 3 \times 8$ Stride: $1 \times 1$ Padding: same
LL->Output-1: Convolutional and Sigmoid activation	Filter size: $31 \times 31 \times 1$ Stride: $1 \times 1$ Padding: same

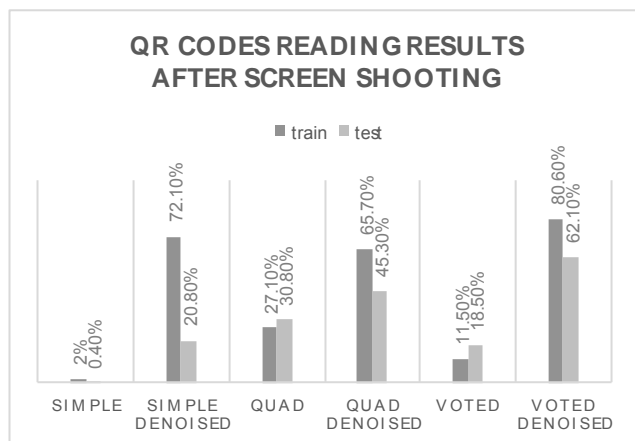
شبکه ی خود شناس CDAE که در کنار U-net برای مقایسه ی نتایج رفع نویز انتخاب شده است از شش لایه ی کانولوشنال با تابع فعال ساز رلو ایجاد شده و در این شبکه نیز در تمامی لایه های کانولوشنال از کرنل های تنظیم کننده L2 با پارامتر  $\lambda$  برابر  $1.0e-3$  استفاده کردیم. در شکل ۹ شمای کلی این شبکه نشان داده شده است.



شکل ۹ ساختار لایه های شبکه ی CDAE [۲۲]

Layer Description	Parameters
Input: gray scale image	Shape(512,512,1)
Input->init: Convolutional and ReLu activation	Filter size: $31 \times 31 \times 3$ Stride: $1 \times 1$ Padding: same
Init->L1: Convolutional and ReLu activation	Filter size: $7 \times 7 \times 64$ Stride: $2 \times 2$ Padding: same
L1->L2: Convolutional and ReLu activation	Filter size: $5 \times 5 \times 128$ Stride: $2 \times 2$ Padding: same
L2->L3: Convolutional and ReLu activation	Filter size: $3 \times 3 \times 192$ Stride: $2 \times 2$ Padding: same
L3->L4: Convolutional and ReLu activation	Filter size: $3 \times 3 \times 256$ Stride: $2 \times 2$ Padding: same
L4->L5: Convolutional and ReLu activation	Filter size: $3 \times 3 \times 384$ Stride: $2 \times 2$ Padding: same
L5->L6: Convolutional and ReLu activation	Filter size: $3 \times 3 \times 512$ Stride: $2 \times 2$ Padding: same
L6->Lx1: Convolutional and ReLu activation	Filter size: $3 \times 3 \times 768$ Stride: $1 \times 1$ Padding: valid
Lx1->Lx1: Convolutional and ReLu activation	Filter size: $3 \times 3 \times 1024$ Stride: $1 \times 1$ Padding: valid
Lx1->Lx2: Convolutional and ReLu activation	Filter size: $3 \times 3 \times 2048$ Stride: $1 \times 1$ Padding: valid
Lx3: Deconvolutional and ReLu activation	Filter size: $3 \times 3 \times 1024$ Stride: $1 \times 1$ Padding: valid
Lx4: Concatenate Lx3 , Lx1	-
Lx4->Last4: Convolutional and ReLu activation (MCNN Layer)	Filter size: $3 \times 3 \times 8$ Stride: $1 \times 1$ Padding: same
Last4->Output-8: Convolutional and Sigmoid activation (MCNN Layer)	Filter size: $3 \times 3 \times 1$ Stride: $1 \times 1$ Padding: same
Lx4->Lx5: Deconvolutional and ReLu activation	Filter size: $3 \times 3 \times 768$ Stride: $1 \times 1$ Padding: valid
Lx5->L7: Deconvolutional and ReLu activation	Filter size: $3 \times 3 \times 512$ Stride: $1 \times 1$ Padding: valid
L9: Concatenate L7 , L6	-
L9->Last8: Convolutional and ReLu activation (MCNN Layer)	Filter size: $3 \times 3 \times 8$ Stride: $1 \times 1$ Padding: same
Last8->Output-7: Convolutional and Sigmoid activation (MCNN Layer)	Filter size: $3 \times 3 \times 1$ Stride: $1 \times 1$ Padding: same
L9->L10: Deconvolutional and ReLu activation	Filter size: $3 \times 3 \times 384$ Stride: $2 \times 2$ Padding: same
L12: Concatenate L10 , L5	-
L12->Last16: Convolutional and ReLu activation (MCNN Layer)	Filter size: $3 \times 3 \times 8$ Stride: $1 \times 1$ Padding: same
Last16->Output-6: Convolutional and Sigmoid activation (MCNN Layer)	Filter size: $3 \times 3 \times 1$ Stride: $1 \times 1$ Padding: same
L12->L13: Deconvolutional and ReLu activation	Filter size: $3 \times 3 \times 256$ Stride: $2 \times 2$ Padding: same
L15: Concatenate L13 , L4	-
L15->Last32: Convolutional and ReLu activation (MCNN Layer)	Filter size: $3 \times 3 \times 8$ Stride: $1 \times 1$ Padding: same
Last32->Output-5: Convolutional and Sigmoid activation (MCNN Layer)	Filter size: $3 \times 3 \times 1$ Stride: $1 \times 1$ Padding: same

## ۵ تنظیمات آموزش



شکل ۱۰ نمودار درصد خوانایی کیو آر های استخراج شده و دینویز شده با شبکه‌ی MCNN برای تمامی رویکرد ها بعد از فرایند اسکرین-دوربین

نتایج شکل ۱۰ شامل استخراج تمامی چهار سطح QR موجود در مجموعه داده بعد از فرایند اسکرین-دوربین می‌باشد. همانطور که مشاهده می‌کنید در تمامی موارد بعد از رفع نویز با شبکه‌ی عصبی نتایج پیشرفت چشمگیری داشته و در بین آن‌ها نتایج رویکرد استخراج رای گیری شده بعد از رفع نویز از تمامی نتایج، بالاتر است. نتایج ذکر شده با آزمایش بر روی داده‌های جمع‌آوری شده در شرایط تصویر برداری مختلف با دوربین‌ها و از روی نمایشگرهای مختلف جمع‌آوری شده‌اند.

جدول ۲ مقایسه‌ی نتایج شبکه‌ی MCNN با دو شبکه‌ی U-net و CDAE

Trained model	Readability (%) Train	Readability (%) Test
MCNN simple	72.1	20.8
U_net simple	32.2	17.4
CDAE simple	2	0.4
MCNN quad	65.7	45.3
U_net quad	46.8	32.3
CDAE quad	28.2	30.8
MCNN voted	80.6	62.1
U_net voted	68.3	42.4
CDAE voted	22.3	19.6

جمع‌آوری و استخراج دیتا ست QRها در سیستم عامل مک نسخه‌ی ۶۴ بیتی با ۱۶ گیگ رم و پردازنده‌ی 6-core 2.6 GHz intel Core i7 و الگوریتم نهان‌نگاری و با استخراج با نرم‌افزار متلب ورژن R2019b پیاده‌سازی شده است. پیاده‌سازی شبکه‌ی عصبی خودشناس با زبان پایتون نسخه‌ی ۳,۷ انجام گرفته است. همچنین برای افزایش سرعت در آموزش شبکه‌ی عصبی و استفاده از پردازنده‌ی گرافیکی برای گرفتن نتایج سریع‌تر و بهتر، از سرویس ابری گوگل تحت عنوان گوگل کلب<sup>۱</sup> که بر پایه‌ی محیط ژوپیتل نوت بوک<sup>۲</sup> می‌باشد، استفاده شده است.

در سرویس ابری گوگل کلب یک پردازنده‌ی گرافیکی NVIDIA Tesla T4 12GB (در گذشته یک پردازنده‌ی گرافیکی Tesla K80) به صورت رایگان در اختیار هر کاربر قرار داده می‌شود که می‌تواند ۱۲ ساعت به طور مداوم مورد استفاده قرار گیرد و کاربر می‌تواند جهت توسعه و اجرای برنامه‌های کاربردی و تحقیقاتی در زمینه یادگیری عمیق، با استفاده از کتابخانه‌های رایج پای تورچ<sup>۳</sup>، تنسورفلو<sup>۴</sup>، کراس<sup>۵</sup> و اپین سی<sup>۶</sup> از آن استفاده نماید، ما در این این پژوهش برای پیاده‌سازی کدهای مربوط به شبکه‌ی عصبی عمیق MCNN از کتابخانه‌های تنسورفلو و کراس و اپین سی وی استفاده نموده ایم.

برای آموزش هر رویکرد استخراج برای هر کدام از ۳ شبکه مجزا، با ساختارهای U-net، CDAE و MCNN، بهینه‌سازی آدام<sup>۷</sup> [۳۳]، ۱۲۰ گام آموزشی<sup>۸</sup> و با اندازه‌ی دسته‌ی ۸ در نظر گرفته شده است.

## ۶ نتایج آزمایش

در این بخش به ارزیابی میزان خوانایی QRهای استخراج شده با رویکردهای مختلف می‌پردازیم. در ابتدا به بررسی‌ی نتایج میزان خوانایی QRهای رفع نویز شده با شبکه‌ی MCNN پرداخته و که در شکل ۱۰ نمودار آن را مشاهده می‌کنید. در ادامه نتایج مقایسه‌ی عملکرد شبکه‌ی MCNN از نظر خوانایی کدهای QR در مقابل دو شبکه‌ی استاندارد U-net و CDAE در جدول ۲ گزارش داده شده است.

<sup>1</sup> Google Colab (<https://colab.research.google.com>)

<sup>2</sup> Jupyter Notebook

<sup>3</sup> PyTorch

<sup>4</sup> TensorFlow

<sup>5</sup> keras

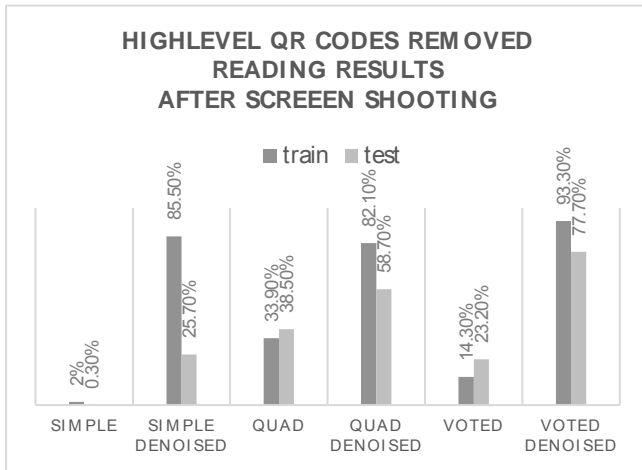
<sup>6</sup> OpenCV

<sup>7</sup> Adam optimizer

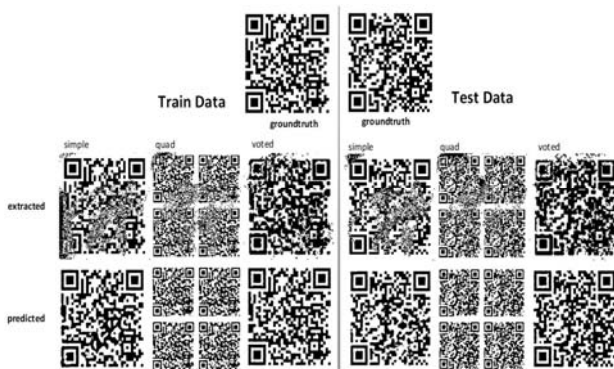
<sup>8</sup> epoch

<sup>9</sup> Batch size





شکل ۱۱ نمودار درصد خوانایی QR های استخراج شده برای تمامی رویکردها بعد از فرایند اسکرین-دوربین بدون احتساب کیو آر های سطح ۴



شکل ۱۲ نتایج بدست آمده روی دو نمونه از داده های تست و آموزش، برای هر سه رویکرد درون نگاری ساده، تکرار چهار گانه و استخراج رای گیری شده. ردیف بالا تصویر: استخراج شده. ردیف پایین: تصاویر تخمین زده شده ی خوانا به وسیله ی شبکه ی عصبی MCNN

## ۷ نتیجه گیری

در مقاله ی پیش رو یک روش کاربردی مبتنی بر شبکه عصبی کانولوشن عمیق چند وضوحه برای حل مسئله ی معکوس رفع نویز از QR های ناخوانا و داراری اعوجاج عمیق مطرح شد. قابلیت عمومی رفع نویز شبکه ی مطرح شده با استفاده از استخراج QR های درون نگاری شده در تصاویر با روش فوریه ی چهارگانه و با سه رویکرد ساده، تکرار چهارگانه و رای گیری شده ارزیابی شده تا نتایج سه شبکه ی آموزش دیده از منظر مجموعه-داده ی غیر شبیه سازی شده، بسیار قابل اعتماد تر باشند. همگرا شدن سریع و روان با تطبیق لایه های خروجی اضافی با ویژگی های فرکانس پایین مربوطه، کاهش تنظیم های بیش از حد پارامتر های شبکه ی عصبی عمیق، از مزایای شبکه ی آموزش دیده می-باشد و لازم به ذکر است که خروجی بسیار قابل قبولی بر روی داده های جمع آوری شده، ارائه شده است و یکی از مسائل معکوس چالش بر انگیز در دنیای واقعی را به خوبی حل کرده است. انتظار

در جدول ۲ برتری نتایج شبکه ی MCNN به وضوح نمایان است، نتایج اخذ شده از شبکه ی MCNN حتی نسبت به شبکه ی U-net که پایه ی ساختاری MCNN را تشکیل می-دهد بسیار بهتر است و این برتری به دلیل اضافه شدن لایه های پنهان چند وضوحه می-باشد.

با بررسی نتایج دریافتیم که سیستم رفع نویز نیاز به آموزش بیشتری در مورد کدهای QR سطح چهار دارد. چرا که به دلیل تراکم زیاد این سطح از QR ها، در بسیاری از مواقع خود QR به عنوان نویز شناخته می-شود و شبکه ی عصبی قادر به بازسازی آن ها نیست. دو دلیل برای این اتفاق وجود دارد. اول این که به نسبت جزئیات این نوع از QR ها تعداد داده ی بیشتری برای آموزش بازیابی و رفع نویز آن ها نیاز است و دوم این که آموزش هر چهار سطح با هم می-تواند بر عملکرد تفکیک پذیر شبکه ی رفع نویز تاثیر منفی بگذارد. یکی از راهکار های بالاتر بردن دقت برای اینگونه داده های پیچیده، آموزش یک شبکه ی مستقل با تعداد بالای داده تنها برای این نوع از QR و استفاده ی آن به صورت موازی در کنار شبکه های دیگر است که در مسیر پژوهش های آینده ی ما قرار خواهد داشت.

برای اطمینان در عملکرد شبکه بدون وجود سطح پیچیدگی چهار، تمامی موارد درون نگاری QR های سطح چهار را حذف کرده و بار دیگر میزان درصدها را بررسی کردیم. مشاهده کردیم که درصد خوانایی همه ی موارد بالاتر رفته و بهترین آن مربوط به رویکرد استخراج رای گیری شده همراه با رفع نویز شبکه ی عصبی بود که از همانطور که در شکل ۹ مشاهده میشود درصد های ۸۰٫۶ برای داده های آموزش و ۶۲٫۱ برای داده های تست به درصد های ۹۳٫۳ و ۷۷٫۷ ارتقا یافت.

نکته ی حائز اهمیت دیگری که وجود دارد بحث انواع الگوریتم برای خوانش QR می-باشد. همانطور که پیش تر اشاره شد در این مقاله از کتابخانه ی pyzbar مبتنی بر زبان پایتون استفاده شد اما در طی آزمایش مشاهده شد که الگوریتم های موجود بر روی گوشی های همراه، بسیاری از مواردی که بعد از رفع نویز با شبکه ی عصبی توسط کتابخانه ی فوق قادر به خوانش نبود را به راحتی می-خوانند و می-توان درصد های بالاتری را با جایگزین کردن QR خوان های درون تلفن های همراه بدست آورد اما برای جلوگیری از اتلاف وقت، در این مقاله نتایج بر اساس کتابخانه نامبرده اعلام شده است.

در شکل ۱۲ نمونه ای از خروجی رفع نویز شده ی شبکه را برای هر سه رویکرد ساده، تکرار چهارگانه و استخراج رای گیری شده از تکرار چهارگانه را مشاهده می-نمایید. همانطور که مشخص است شبکه قدرت بالایی در تشخیص و پیشبینی مناطق از بین رفته ی QR دارد.

- [۱۳] Z. Yang, Z. Dai, Y. Yang, J. Carbonell, R. Salakhutdinov, and Q. V. Le, "Xlnet: Generalized autoregressive pretraining for language understanding," *arXiv preprint arXiv:1906.08237*, 2019.
- [۱۴] A. Brock, J. Donahue, and K. Simonyan, "Large scale GAN training for high fidelity natural image synthesis," *arXiv preprint arXiv:1809.11096*, 2018.
- [۱۵] R. Luo, F. Tian, T. Qin, E. Chen, and T.-Y. Liu, "Neural architecture optimization," *arXiv preprint arXiv:1808.07233*, 2018.
- [۱۶] H. Jin, Q. Song, and X. Hu, "Auto-keras: An efficient neural architecture search system," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2019, pp. 1946-1956.
- [۱۷] L. N. Smith, "A disciplined approach to neural network hyper-parameters: Part 1--learning rate, batch size, momentum, and weight decay," *arXiv preprint arXiv:1803.09820*, 2018.
- [۱۸] X.-y. Wang, C.-p. Wang, H.-y. Yang, and P.-p. Niu, "A robust blind color image watermarking in quaternion Fourier transform domain," *Journal of Systems and Software*, vol. 86, no. 2, pp. 255-277, 2013.
- [۱۹] S.-C. Pei, J.-J. Ding, and J.-H. Chang, "Efficient implementation of quaternion Fourier transform, convolution, and correlation by 2-D complex FFT," *IEEE Transactions on Signal Processing*, vol. 49, no. 11, pp. 2783-2797, 2001.
- [۲۰] M. Monfared, A. Koochari, and R. Monshianmotlagh, "QR-DN1. 0: A new distorted and noisy QRs dataset," *Data in Brief*, p. 107605, 2021.
- [۲۱] P. Vincent, H. Larochelle, Y. Bengio, and P. Manzagol, "Proceedings of the 25th international conference on Machine learning," 2008.
- [۲۲] D. Lee, S. Choi, and H.-J. Kim, "Performance evaluation of image denoising developed using convolutional denoising autoencoders in chest radiography," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 884, pp. 97-104, 2018.
- [۲۳] Z.-Q. J. Xu, Y. Zhang, and Y. Xiao, "Training behavior of deep neural network in frequency domain," in *International Conference on Neural Information Processing*, 2019: Springer, pp. 264-274.
- [۲۴] N. Rahaman *et al.*, "On the spectral bias of neural networks," in *International Conference on Machine Learning*, 2019: PMLR, pp. 5301-5310.
- [۲۵] K. Stüben and U. Trottenberg, "Multigrid methods: Fundamental algorithms, model problem analysis and applications," *Multigrid methods*, pp. 1-176, 1982.
- [۲۶] F. Wang, A. Eljarrat, J. Müller, T. R. Henninen, R. Erni, and C. T. Koch, "Multi-resolution convolutional می‌رود که با استفاده از ساختارهای جدید تر و مجموعه داده‌ی آموزشی بیشتر و متنوع تر از نظر اعوجاج‌ها و همچنین آموزش انحصاری هر شبکه‌ی عصبی برای هر سطح از QR و استفاده‌ی موازی از این شبکه‌ها، بتوان گام بزرگی در جهت رفع اعوجاج و نویزهای بارکد‌های حساس مانند QR برداشت.
- ## مراجع
- [۱] T. J. Soon, "QR code," *Synthesis Journal*, vol. 2008, pp. 59-78, 2008.
- [۲] O. Russakovsky *et al.*, "Imagenet large scale visual recognition challenge," *International journal of computer vision*, vol. 115, no. 3, pp. 211-252, 2015.
- [۳] Y. Li, X. Gao, and H. Ji, "A 3D wavelet based spatial-temporal approach for video watermarking," in *Proceedings Fifth International Conference on Computational Intelligence and Multimedia Applications. ICCIMA 2003*, 2003: IEEE, pp. 260-265.
- [۴] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, pp. 436-444, 2015.
- [۵] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE transactions on pattern analysis and machine intelligence*, vol. 38, no. 2, pp. 295-307, 2015.
- [۶] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising," *IEEE transactions on image processing*, vol. 26, no. 7, pp. 3142-3155, 2017.
- [۷] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1125-1134.
- [۸] Y. D. Hezaveh, L. P. Levasseur, and P. J. Marshall, "Fast automated analysis of strong gravitational lenses with convolutional neural networks," *Nature*, vol. 548, no. 7669, pp. 555-557, 2017.
- [۹] A. Sinha, J. Lee, S. Li, and G. Barbastathis, "Lensless computational imaging through deep learning," *Optica*, vol. 4, no. 9, pp. 1117-1125, 2017.
- [۱۰] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural networks*, vol. 2, no. 5, pp. 359-366, 1989.
- [۱۱] G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Mathematics of control, signals and systems*, vol. 2, no. 4, pp. 303-314, 1989.
- [۱۲] Z. Lu, H. Pu, F. Wang, Z. Hu, and L. Wang, "The expressive power of neural networks: A view from the width," *arXiv preprint arXiv:1709.02540*, 2017.



میلاذ منفرد مدرک کارشناسی ارشد خود را در رشته ی کامپیوتر گرایش هوش مصنوعی در سال ۱۴۰۰ دریافت کرده و زمینه های پژوهشی مورد علاقه ی وی پردازش تصویر، بینایی ماشین، یادگیری عمیق و امنیت است.



عباس کوچاری دکتری خود را در سال ۱۳۹۱ در رشته ی کامپیوتر گرایش هوش مصنوعی اخذ نموده و هم اکنون استادیار و عضو هیئت عملی دانشگاه آزاد واحد علوم و تحقیقات می باشد. زمینه های پژوهشی مورد علاقه ی وی پردازش تصویر، بینایی ماشین، پردازش گفتار و زبان های طبیعی و یادگیری عمیق است.

- neural networks for inverse problems," *Scientific reports*, vol. 10, no. 1, pp. 1-11, 2020.
- [۲۷] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *science*, vol. 313, no. 5786, pp. 504-507, 2006.
- [۲۸] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, 2015: Springer, pp. 234-241 .
- [۲۹] T. Karras, T. Aila, S. Laine, and J. Lehtinen, "Progressive growing of gans for improved quality, stability, and variation," *arXiv preprint arXiv:1710.10196*, 2017.
- [۳۰] S. Nah, T. Hyun Kim, and K. Mu Lee, "Deep multi-scale convolutional neural network for dynamic scene deblurring," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 3883-3891 .
- [۳۱] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *International conference on machine learning*, 2015: PMLR, pp. 448-456 .
- [۳۲] Y. Wu and K. He, "Group normalization," in *Proceedings of the European conference on computer vision (ECCV)*, 2018, pp. 3-19 .
- [۳۳] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.