

## بهبود کیفیت تصاویر دارای اعوجاج فوکوس با استفاده از شبکه‌های عصبی کانولوشنی

فاطمه فلاحی، زهرا فلاحی، محمدرضا جعفر فرد<sup>\*</sup>، بابک زارع

دانشکده فیزیک، دانشگاه علم و صنعت ایران، ایران

### چکیده

### اطلاعات مقاله

اعوجاج فوکوس یکی از عوامل اصلی کاهش کیفیت تصاویر میکروسکوپی است و می‌تواند موجب کاهش وضوح، از دست رفتن جزئیات و افت دقت در تحلیل تصاویر شود. از این رو، توسعه روش‌های مؤثر برای اصلاح این اعوجاج از اهمیت بالایی برخوردار است. در این پژوهش، یک شبکه عصبی رمزگذار خودکار کانولوشنی متشکل از لایه‌های کانولوشن، گلوگاه فشرده و لایه‌های کانولوشن انتقالی برای اصلاح اعوجاج فوکوس در تصاویر میکروسکوپی دیجیتال طراحی و آموزش داده شد. بدین منظور، یک مجموعه داده شامل تصاویر مرجع و تصاویر دارای اعوجاج فوکوس مبتنی بر مدل پراش فرنل تهیه و برای آموزش و ارزیابی شبکه مورد استفاده قرار گرفت. عملکرد مدل با استفاده از معیارهای میانگین مربع خطا (MSE)، نسبت سیگنال به نویز پیک (PSNR) و شاخص تشابه ساختاری (SSIM) ارزیابی شد. نتایج به دست آمده شامل MSE برابر با ۰.۰۰۸/۰، PSNR برابر با ۹۷/۳۰ دسی‌بل و SSIM برابر با ۹۸۳۹/۰ است که بیانگر بازسازی دقیق و شباهت ساختاری بسیار بالای تصاویر بازسازی شده نسبت به تصاویر مرجع است. این نتایج نشان می‌دهد که مدل پیشنهادی می‌تواند به عنوان رویکردی مؤثر برای حذف اعوجاج فوکوس، بهبود وضوح تصاویر میکروسکوپی و ارتقای کیفیت بازسازی تصاویر مورد استفاده قرار گیرد.

تاریخ دریافت: ۱۴۰۴/۱۲/۰۵

تاریخ پذیرش: ۱۴۰۵/۰۶/۱۹

تاریخ چاپ: ۱۴۰۵/۰۷/۰۸

شاپای چاپی: ۴۹۲۱-۲۵۸۸

شاپای الکترونیکی: ۴۹۳۸-۲۵۸۸

\* نویسنده مسئول

[mrjafarfard@iust.ac.ir](mailto:mrjafarfard@iust.ac.ir)

DOI:

۱۰.۲۲۰۳۴/jmrph.۲۰۲۶.۱۲۲۹۲۳.۰

**واژگان کلیدی:** فوکوس، پراش فرنل، یادگیری عمیق، شبکه عصبی کانولوشنی.

### مقدمه

شدید کیفیت تصویر می‌شود، بلکه در بسیاری از کاربردها هزینه‌بر و مشکل ساز است. اعوجاج فوکوس زمانی رخ می‌دهد که محل تشکیل تصویر دقیقاً منطبق بر صفحه حسگر دوربین نباشد، در نتیجه تصویر حاصل تار یا دارای حلقه‌های پراشی ناخواسته خواهد بود. این اعوجاج می‌تواند وضوح تصویر را به طور قابل توجهی کاهش داده و کاربرد آن را در زمینه‌هایی مانند عکاسی، تصویربرداری پزشکی و تصویربرداری ماهواره‌ای تحت تاثیر قرار دهد. برای رفع این مشکل، تاکنون روش‌های متعددی پیشنهاد شده‌اند که عمدتاً در دو دسته کلی قرار می‌گیرند: تکنیک‌های محاسباتی پردازش تصویر [۱] و روش‌های اپتیکی [۲، ۳]. روش‌های اپتیکی شامل اصلاح لنز، کالیبراسیون دقیق

پیشرفت‌های اخیر در فناوری تصویربرداری دیجیتال باعث شده‌اند که امروزه تصاویر به عنوان داده‌های اولیه، نقش بسیار مهمی در صنایع مختلف ایفا کنند. وضوح تصویر، عاملی کلیدی در کیفیت، تفسیر و کاربرد آن به شمار می‌رود. با این حال، دستیابی به وضوح بالا همواره آسان نیست، چرا که عوامل متعددی نظیر نویزهای اپتیکی، ابیراهی‌های رنگی و کروی، واپیچشها<sup>۱</sup> و نبود فوکوس می‌توانند منجر به تار یا اعوجاج تصویر شوند. از میان این عوامل، عدم فوکوس صحیح اهمیت ویژه‌ای دارد، زیرا نه تنها باعث کاهش

<sup>۱</sup> Distortion

فوکوس است. این رویکرد، ضمن حفظ مبنای فیزیکی فرآیند تشکیل تصویر، امکان یادگیری نگاشت بین تصاویر دارای اعوجاج و تصاویر مرجع را بدون نیاز به مدل‌سازی صریح مسئله معکوس فراهم می‌کند. نتایج حاصل نیز نشان می‌دهد که مدل پیشنهادی قادر است تصاویر دارای اعوجاج فوکوس را با دقت بالا بازسازی کرده و کیفیت آن‌ها را به‌طور مؤثر بهبود بخشد.

## ۱- مبانی نظری

### ۱.۲. فوکوس

مفهوم فوکوس در تصویربرداری، به تنظیم فاصله کانونی لنز دوربین بر اساس فاصله جسم تا دوربین برای دستیابی به تصویری واضح اطلاق می‌شود. در کاربردهایی که نمایش دقیق و تحلیل نمونه اهمیت بالایی دارد، مانند تصویربرداری پزشکی، میکروسکوپی، تصویربرداری نظامی، هواشناسی و سایر حوزه‌های علمی، اعوجاج فوکوس می‌تواند به دلیل ماهیت حیاتی اطلاعات منتقل شده، منجر به نتایج خطرناک یا تفسیر نادرست شود. بنابراین، حصول اطمینان از کیفیت بالای تصاویر و حذف اعوجاج فوکوس، برای دستیابی به نتایج دقیق و قابل اعتماد، از اهمیت بالایی برخوردار است [۱۰، ۱۱].

در زمینه بررسی فوکوس و تاری تصویر، پدیده‌ی پراش فرنل می‌تواند نقش مؤثری ایفا کند، به‌ویژه در شرایطی که نور از دیافراگم‌های کوچک عبور کرده یا با موانعی برخورد کند. در این حالت، عبور نور از دیافراگم باعث تغییر در شدت و فاز موج می‌شود که می‌تواند بر کیفیت فوکوس تأثیرگذار باشد. تحلیل این پدیده با استفاده از مدل پراش فرنل امکان‌پذیر است که به کمک آن می‌توان فاصله دقیق بین جسم، لنز و حسگر تصویر را محاسبه کرد؛ پارامترهایی که بر وضوح تصویر تأثیر می‌گذارند [۱۲].

فوکوس خودکار و استفاده از تجهیزات پیشرفته هستند که گرچه دقت بالایی دارند، اما بسیار پرهزینه و وابسته به تجهیزات خاص‌اند و عموماً پیش از تصویربرداری باید اعمال شوند. از سوی دیگر، روش‌های محاسباتی مبتنی بر ارزیابی فوکوس به‌صورت عددی، با استفاده از معیارهایی مانند گرادیان مطلق<sup>۲</sup>، گرادیان مربع<sup>۳</sup>، فیلتر لاپلاسی<sup>۴</sup>، تابع برنر و ... عمل می‌کنند. به عنوان مثال، در اندازه‌گیری دقیق حجم یک نمونه، وجود حتی مقدار اندکی اعوجاج فوکوس می‌تواند موجب خطای محسوس در نتایج شود. در چنین مواردی، یافتن نقطه فوکوس بهینه با بررسی سطح نمونه و تحلیل همزمان معیارهای فوکوس مانند SGA (الگوریتم گرادیان مربع)<sup>۵</sup> و LFA (الگوریتم فیلتر لاپلاسی)<sup>۶</sup> ضروری است. با این حال، روش‌های محاسباتی معمولاً در زمان تصویربرداری اعمال می‌شوند، فرآیندی زمانبر که همواره دقت لازم را تضمین نمی‌کند. به‌ویژه در شرایطی که امکان تصویربرداری مجدد وجود ندارد، این روش‌ها کارایی محدودی دارند.

در سال‌های اخیر، پیشرفت‌های چشمگیر در حوزه یادگیری عمیق، به‌ویژه شبکه‌های عصبی کانولوشنی، تحول قابل توجهی در زمینه بازسازی و بهبود کیفیت تصاویر ایجاد کرده است. این روش‌ها به دلیل توانایی بالا در استخراج خودکار ویژگی‌ها و مدل‌سازی روابط غیرخطی، در مسائل حذف تاری، اصلاح اعوجاج و بازیابی تصاویر کاربرد گسترده‌ای یافته‌اند [۴-۹]. با وجود این، بخش عمده مطالعات انجام‌شده بر حذف تاری عمومی<sup>۷</sup> یا بهبود کیفیت تصاویر طبیعی متمرکز بوده‌اند و پژوهش‌های کمتری به مسئله اصلاح اعوجاج فوکوس مبتنی بر یک مدل فیزیکی معتبر در تصویربرداری میکروسکوپی پرداخته‌اند.

در این مقاله، مسئله حذف اعوجاج فوکوس با استفاده از یک شبکه عصبی رمزگذار خودکار کانولوشنی مورد بررسی قرار گرفته است. برای این منظور، ابتدا مجموعه‌ای از تصاویر واضح تهیه و با استفاده از مدل پراش فرنل، نسخه‌های دارای اعوجاج فوکوس از آن‌ها تولید شد. سپس شبکه پیشنهادی با استفاده از این داده‌ها آموزش داده شد. وجه تمایز این پژوهش در ترکیب مدل فیزیکی پراش فرنل برای تولید داده‌های آموزشی با یک شبکه رمزگذار خودکار کانولوشنی به‌منظور بازسازی مستقیم تصاویر دارای اعوجاج

<sup>۲</sup> Absolute Gradient

<sup>۳</sup> Square Gradient

<sup>۴</sup> Laplacian Filter

<sup>۵</sup> Squared Gradient Algorithm

<sup>۶</sup> Laplace Filter Algorithm

<sup>۷</sup> Image Deblurring

### ۲.۲. پراش فرنل

MSE به صورت زیر تعریف می شود:

$$MSE = \frac{1}{MN} \sum_{X=0}^{M-1} \sum_{Y=0}^{N-1} [X'(X,Y) - X(X,Y)]^2$$

که در آن  $X(X,Y)$  تصویر مرجع و  $X'(X,Y)$  تصویر بازسازی شده است. با رابطه زیر داده می شود:

$$PSNR = 10 \times \log_{10} \frac{MAX^2}{MSE}$$

که در آن  $MAX$  بیشینه مقدار شدت تصویر (برای تصاویر ۸ بیتی برابر ۲۵۵) است و نهایتاً  $SSIM$  با عبارت زیر تعریف می شود:

$$SSIM(X,Y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}$$

که در آن  $\mu_x$  و  $\mu_y$  میانگین شدت،  $\sigma_x^2$  و  $\sigma_y^2$  واریانس، و  $\sigma_{xy}$  کوواریانس دو تصویر بوده و  $C_1$  و  $C_2$  ضرایب تثبیت کننده هستند.

خطای میانگین مربعات تفاوت بین مقدار اصلی و مقدار بازسازی شده را نشان می دهد و هرچه مقدار آن کمتر باشد، دقت بازسازی بیشتر است. نسبت سیگنال به نویز پیک نسبت بین بیشینه توان سیگنال به توان نویز است. مقدار بالاتر  $PSNR$  نشان دهنده کیفیت بهتر بازسازی است.  $SSIM$  میزان شباهت ساختاری بین دو تصویر را می سنجد و درک بصری انسان را بهتر نشان می دهد.

### ۳. شبیه سازی

#### ۳.۱. شبکه عصبی مورد استفاده

برای رفع اعوجاج فوکوس، از یک شبکه عصبی رمزگذار خودکار کانولوشنی<sup>۱۲</sup> استفاده شده است. این نوع شبکه ها با بهره گیری از ویژگی فشرده سازی و بازسازی، قادر به بازیابی جزئیات تصویر هستند. همانطور که در شکل (۱)

در بحث فوکوس خودکار، مدل پراش فرنل کاربرد بالایی دارد، زیرا در ناحیه میدان نزدیک، که جبهه موج هنوز به حالت مسطح نرسیده، امکان تحلیل دقیق تری از فاز موج فراهم می شود؛ مسئله ای که برای فوکوس دقیق ضروری است. وقتی یک جسم کدر بین یک منبع نقطه ای نور و پرده تصویر قرار گیرد، الگوی سایه ای پیچیده با نواحی روشن و تاریک تشکیل می شود. این الگو که به وضوح قابل مشاهده است، بازتابی از پراش فرنل یا "میدان نزدیک"<sup>۱۳</sup> است [۱۳، ۱۴].

به طور خاص، زمانی که امواج تخت حاصل از منبع نوری از روزنه ای کوچک عبور می کنند، تصویری از روزنه روی پرده تشکیل می شود. با افزایش فاصله پرده از مانع، تصویر روزنه همچنان قابل تشخیص است ولی فریزها بیشتر نمایان می شوند. این رفتار ویژگی بارز میدان نزدیک است.

میدان الکتریکی در نقطه  $(x,y,z)$  طبق رابطه ی زیر از مدل پراش فرنل محاسبه می شود:

$$E(x,y,z) = \frac{e^{ikz}}{i\lambda z} \iint E(x',y',0) e^{\frac{ik}{2z}[(x-x')^2 + (y-y')^2]} dx' dy' \quad (1)$$

در این رابطه،  $E(x,y,z)$  میدان در صفحه خروجی (تصویر)،  $E(x',y',0)$  میدان ورودی،  $\lambda$  طول موج نور،  $k$  عدد موج و  $z$  فاصله روزنه تا دوربین است.

#### ۳.۲. معیارهای ارزیابی کیفیت تصویر

ارزیابی کیفیت تصویر به معنای تخمین عددی میزان وضوح و دقت تصویر است و معمولاً برای سنجش عملکرد الگوریتم های پردازش تصویر به کار می رود. این ارزیابی می تواند به دو روش ذهنی و عینی انجام شود. در روش ذهنی از قضاوت انسان استفاده می شود، در حالی که روش عینی متکی بر الگوریتم های محاسباتی است که در نبود تصویر مرجع نیز قابل استفاده اند. در سال های اخیر، به دلیل دقت و تکرارپذیری بالا، ارزیابی های عینی بیشتر مورد توجه قرار گرفته اند. از جمله مهم ترین معیارهای ارزیابی کیفیت تصویر عبارتند از: میانگین مربع خطا<sup>۱۴</sup> ( $MSE$ )، نسبت سیگنال به نویز پیک<sup>۱۵</sup> ( $PSNR$ ) و شاخص تشابه ساختاری<sup>۱۱</sup> ( $SSIM$ ) [۱۵].

<sup>۱۳</sup> Near Field

<sup>۱۴</sup> Mean Squared Error (MSE)

<sup>۱۵</sup> Peak To Noise Signal Ratio (PSNR)

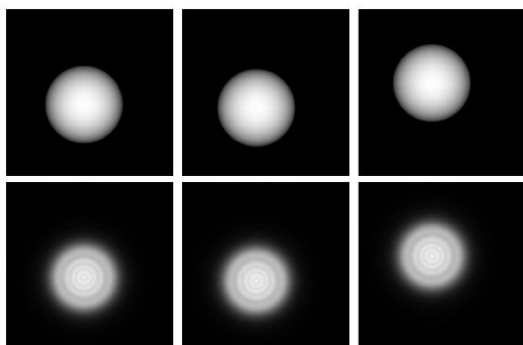
<sup>۱۱</sup> Structural similarity index measure (SSIM)

<sup>۱۲</sup> Convolutional Autoencoder

شکل ۱. نمای شماتیک شبکه عصبی رمزگذار خودکار، شامل بخش رمزگذار برای فشرده‌سازی داده‌های ورودی، گلوگاه به‌عنوان رابط بین رمزگذار و رمزگشا، و بخش رمزگشا برای بازسازی تصویر اولیه [۱۹].

### ۲.۳. مجموعه داده مورد استفاده

برای آموزش شبکه، مجموعه داده‌ای شامل ۱۰۰۰ تصویر  $256 \times 256$  پیکسلی از کره‌هایی با شعاع ثابت و مراکز تصادفی تولید شده است. سپس با استفاده از رابطه پراش فرنل (۱)، نسخه‌های دارای اعوجاج فوکوس از این تصاویر ایجاد شد. برای این منظور، طول موج نور برابر با  $532\text{nm}$  و فاصله روزنه تا دوربین برابر با  $0.2\text{mm}$  در نظر گرفته شده است. در مجموع، ۲۰۰۰ تصویر شامل نسخه‌های واضح و با اعوجاج فوکوس تولید شده است. نمونه‌ای از داده‌های تولید شده در شکل (۳) نشان داده شده است. ۲۰٪ از این داده‌ها به منظور اعتبارسنجی مدل و ارزیابی عملکرد در طول فرآیند آموزش اختصاص داده شده است.



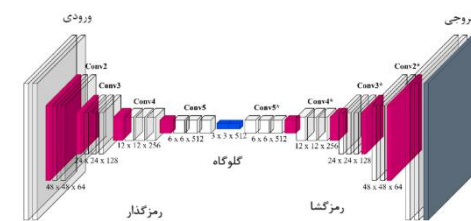
شکل ۲. نمونه‌ای از تصاویر بدون اعوجاج فوکوس (ردیف بالا) و تصاویر دارای اعوجاج فوکوس (ردیف پایین) که با اعمال رابطه‌ی پراش فرنل بر مجموعه تصاویر اصلی تولید شده‌اند.

فرآیند آموزش به‌صورت نظارت نشده<sup>۱۷</sup> و طی ۱۰۰ دوره آموزش<sup>۱۸</sup> انجام شده و از تابع هزینه<sup>۱۹</sup> MSE به‌همراه بهینه‌ساز<sup>۲۰</sup> Adam استفاده شده است.

نشان داده شده، ساختار این شبکه عصبی از سه بخش اصلی رمزگذار<sup>۱۳</sup>، گلوگاه<sup>۱۴</sup> و رمزگشا<sup>۱۵</sup> تشکیل شده است [۱۶ و ۱۷ و ۱۸].

در بخش رمزگذار، سه لایه کانولوشن متوالی با تعداد فیلترهای ۱۶، ۳۲ و ۶۴ به‌ترتیب استفاده شده‌اند. اندازه کرنل‌ها در تمامی لایه‌ها برابر  $3 \times 3$  در نظر گرفته شده و تابع فعال‌سازی ReLU برای افزایش غیرخطی بودن مدل به‌کار رفته است. پس از هر لایه کانولوشن، عملیات MaxPooling برای کاهش ابعاد ویژگی‌ها و استخراج نمایش فشرده تصویر اعمال شده است. همچنین برای جلوگیری از بیش‌برازش، از لایه Dropout با نرخ ۰.۱ استفاده شده است. در بخش گلوگاه، یک لایه کاملاً متصل به‌عنوان نمایش فشرده نهایی با ابعاد کاهش‌یافته به‌کار گرفته شده است که نقش نگاشت ویژگی‌های استخراج‌شده را بر عهده دارد. در بخش رمزگشا، ابتدا عملیات Reshape برای بازسازی ساختار فضایی ویژگی‌ها انجام می‌شود و سپس سه لایه کانولوشن ترنسپوز (Conv2DTranspose) با تعداد فیلترهای متناظر ۱۶، ۸ و ۱ استفاده شده است. برای افزایش ابعاد ویژگی‌ها، از لایه UpSampling بهره گرفته شده و در نهایت یک لایه کانولوشن با یک فیلتر و تابع فعال‌سازی خطی برای بازسازی تصویر خروجی استفاده شده است.

تابع هزینه مورد استفاده در این شبکه، MSE بوده و بهینه‌سازی پارامترها با استفاده از الگوریتم Adam انجام شده است. نرخ یادگیری برابر ۰.۰۰۱ و اندازه دسته<sup>۱۶</sup> برابر ۳۲ در نظر گرفته شده است. همچنین آموزش شبکه به مدت ۱۰۰ epoch انجام شده است. برای ارزیابی عملکرد مدل، از مجموعه اعتبارسنجی در طول فرآیند آموزش استفاده شده است.



<sup>۱۷</sup> Unsupervised Learning

<sup>۱۸</sup> Epoch

<sup>۱۹</sup> Loss Function

<sup>۲۰</sup> Optimizer

<sup>۱۳</sup> Encoder

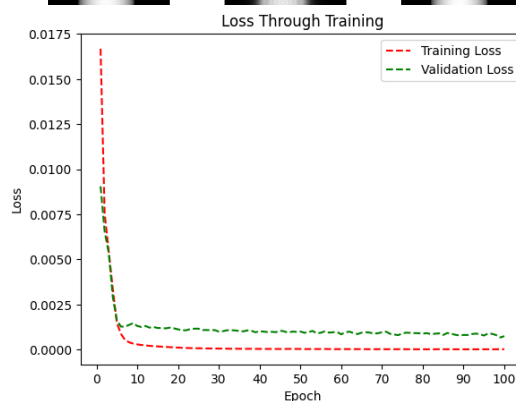
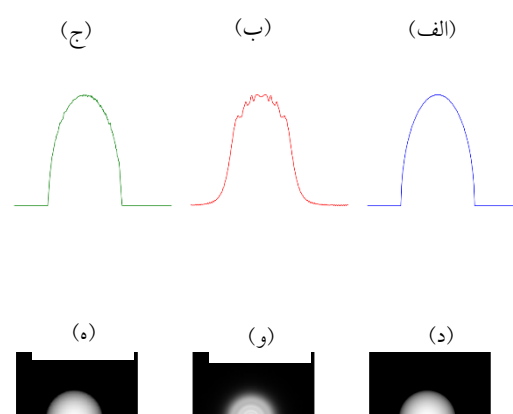
<sup>۱۴</sup> Latent Space

<sup>۱۵</sup> Decoder

<sup>۱۶</sup> Batch size

## ۴. نتایج

به منظور ارزیابی عملکرد شبکه عصبی رمزگذار خودکار کانولوشنی در رفع اعوجاج فوکوس، از مجموعه داده‌های تولید شده برای آموزش و اعتبارسنجی مدل استفاده شد. از آنجا که مسئله حاضر یک مسئله بازسازی تصویر (رگرسیون) است، ارزیابی عملکرد مدل بر اساس معیارهای کمی استاندارد در بازسازی تصویر شامل MSE، PSNR و SSIM انجام گرفت. نتایج حاصل نشان می‌دهد که شبکه آموزش دیده قادر است ساختار و جزئیات تصاویر مرجع را با دقت بسیار بالا بازسازی کند. همانگونه که در شکل (۳) مشاهده می‌شود، تصویر خروجی شبکه از نظر ساختار شدت، لبه‌ها و جزئیات فضایی شباهت قابل توجهی با تصویر مرجع دارد.



(ی)

شکل ۳. نتایج عملکرد شبکه عصبی در رفع اعوجاج فوکوس. (الف) تصویر

مرجع بدون اعوجاج، (ب) تصویر دارای اعوجاج حاصل از اعمال رابطه پراش فرنل، (ج) خروجی شبکه پس از اصلاح اعوجاج فوکوس، (د-ه) نمودارهای برش عرضی در راستای افقی به ترتیب از تصاویر (الف)، (ب) و (ج) جهت مقایسه پروفایل شدت در هر مرحله، (ی) نمودار میزان خطا در طی فرآیند آموزش شبکه برای ۱۰۰۰ داده آموزشی.

به منظور بررسی روند همگرایی شبکه در طول فرآیند آموزش، تغییرات تابع هزینه برای مجموعه‌های آموزش و اعتبارسنجی برحسب تعداد دوره‌های آموزش در شکل (۳) ارائه شده است. همان‌گونه که مشاهده می‌شود، مقدار تابع هزینه در هر دو مجموعه در چند دوره ابتدایی به سرعت کاهش یافته و سپس به مقادیر تقریباً ثابتی همگرا می‌شود. همچنین، نزدیک بودن روند تغییرات خطای آموزش و اعتبارسنجی و عدم افزایش محسوس خطای اعتبارسنجی در طول آموزش، بیانگر پایداری فرآیند یادگیری و عدم بروز بیش‌برازش قابل توجه در مدل است.

مقادیر به دست آمده برای معیارهای ارزیابی نشان می‌دهد که شبکه پیشنهادی به مقدار MSE برابر با ۰/۰۰۰۸، PSNR برابر با ۳۰/۹۷ دسی‌بل و SSIM برابر با ۰/۹۸۳۹ دست‌یافته است. مقدار بسیار کوچک MSE بیانگر اختلاف اندک بین تصویر بازسازی شده و تصویر مرجع است. همچنین مقدار نسبتاً بالای PSNR نشان‌دهنده کیفیت مناسب تصویر بازسازی شده بوده و مقدار SSIM نزدیک به یک نیز بیانگر حفظ موفق ساختارهای فضایی و شباهت ادراکی بالای تصاویر بازسازی شده نسبت به تصاویر مرجع است. این نتایج نشان می‌دهند که شبکه پیشنهادی قادر است اعوجاج فوکوس را به صورت مؤثر حذف کرده و کیفیت تصاویر را به میزان قابل توجهی بهبود دهد.

عملکرد شبکه‌های عصبی به عوامل متعددی از جمله معماری شبکه، پارامترهای آموزش، کیفیت داده‌ها و اندازه مجموعه داده آموزشی وابسته است. به منظور بررسی تأثیر اندازه مجموعه داده، سه مجموعه آموزشی شامل ۱۰۰۰، ۲۰۰۰ و ۳۰۰۰ تصویر مورد استفاده قرار گرفت. در تمامی حالات، سهم داده‌های اعتبارسنجی ۲۰٪ از کل داده‌های آموزشی تعیین شد. سایر پارامترهای شبکه در طول فرآیند یادگیری ثابت نگه داشته شدند تا تأثیر دقیق تغییر اندازه مجموعه داده مشخص شود. نتایج این مطالعه در شکل (۴) و جدول (۱) ارائه شده است.

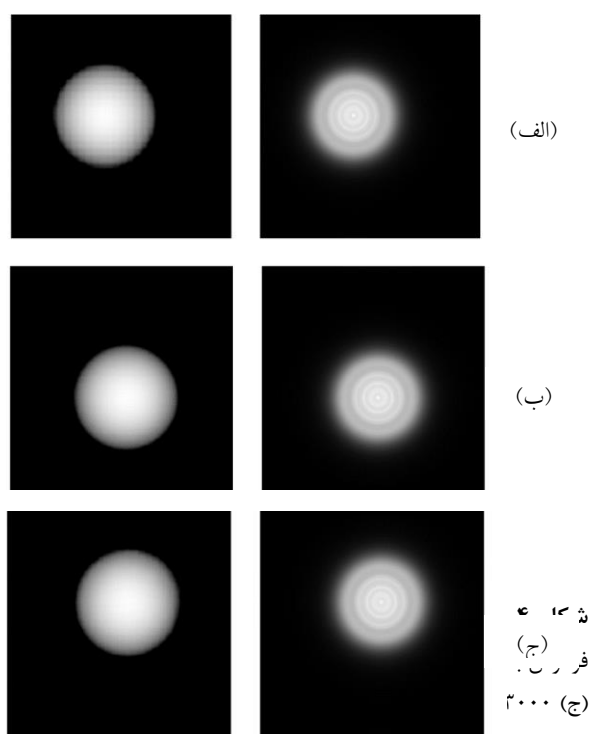
| داده اعتبار سنجی  | ۲۰۰    | ۴۰۰    | ۶۰۰    |
|-------------------|--------|--------|--------|
| MSE               | ۰/۰۰۰۵ | ۰/۰۰۰۸ | ۰/۰۰۰۶ |
| PSNR              | ۳۲/۹۳  | ۳۰/۹۷  | ۳۲/۸۸  |
| SSIM              | ۰/۹۷۸۱ | ۰/۹۸۳۹ | ۰/۹۸۹۶ |
| زمان اجرا (ثانیه) | ۲۸۶۰   | ۴۶۳۲   | ۵۶۳۲   |

#### ۴. نتیجه‌گیری

در این پژوهش، یک شبکه‌ی عصبی کانولوشنی رمزگذار خودکار برای رفع اعوجاج فوکوس طراحی و با استفاده از داده‌های مصنوعی مبتنی بر مدل پراش فرنل آموزش داده شد. فرآیند آموزش طی ۱۰۰ دوره و با استفاده از تابع هزینه میانگین مربعات خطا و بهینه‌ساز Adam انجام گرفت. ارزیابی عملکرد مدل با استفاده از معیارهای کمی استاندارد بازسازی تصویر شامل MSE، PSNR و SSIM انجام شد. نتایج نشان می‌دهند که مدل پیشنهادی به مقادیر MSE برابر ۰/۰۰۰۸، PSNR برابر ۳۰/۹۷ و SSIM برابر ۰/۹۸۳۹ دست‌یافته است. مقدار بالای SSIM بیانگر حفظ ساختارهای فضایی و شباهت ادراکی بسیار بالا بین تصاویر بازسازی‌شده و تصاویر مرجع می‌باشد. بر اساس این نتایج، شبکه طراحی‌شده توانایی بالایی در بازسازی تصاویر و حذف اعوجاج فوکوس از خود نشان می‌دهد. مقایسه عملکرد این روش با رویکردهای مرسوم پردازش تصویر حاکی از آن است که استفاده از یادگیری عمیق می‌تواند راهکاری دقیق‌تر و کارآمدتر برای اصلاح اعوجاج فوکوس در تصاویر میکروسکوپی ارائه دهد.

در این پژوهش، داده‌های آموزشی به‌صورت مصنوعی و مبتنی بر مدل پراش فرنل تولید شدند تا یک چارچوب کنترل‌شده برای ارزیابی اولیه عملکرد شبکه فراهم گردد و اثر عوامل مزاحم موجود در تصاویر واقعی حذف شود. اگرچه این رویکرد امکان تحلیل دقیق عملکرد مدل را فراهم می‌کند، اما می‌تواند محدودیت‌هایی در تعمیم‌پذیری به داده‌های واقعی ایجاد کند. بنابراین، ارزیابی مدل بر روی تصاویر واقعی میکروسکوپی به‌عنوان یکی از مسیرهای توسعه آتی این پژوهش در نظر گرفته می‌شود.

#### منابع



همانطور که در شکل (۴) و جدول (۱) مشاهده می‌شود، افزایش تعداد داده‌های آموزشی الزاماً به بهبود یکنواخت تمامی معیارهای ارزیابی منجر نشده است. به‌طور مشخص، مقدار MSE برای مجموعه ۲۰۰۰ تصویری اندکی بزرگ‌تر از دو حالت دیگر است، در حالی که مقدار SSIM با افزایش تعداد داده‌ها روندی افزایشی نشان می‌دهد و برای مجموعه ۳۰۰۰ تصویری به بیشترین مقدار خود می‌رسد. این رفتار را می‌توان ناشی از افزایش تنوع نمونه‌های آموزشی و ماهیت تصادفی فرآیند آموزش شبکه دانست که موجب می‌شود معیارهای مختلف حساسیت یکسانی نسبت به تغییرات مجموعه داده نداشته باشند.

با وجود این، نتایج نشان می‌دهد که استفاده از مجموعه داده بزرگ‌تر موجب حفظ بهتر ساختارهای تصویر و بهبود قابلیت تعمیم مدل شده است، هرچند این بهبود لزوماً به‌صورت یکنواخت در تمامی معیارهای کمی مشاهده نمی‌شود. از سوی دیگر، افزایش حجم داده‌های آموزشی با افزایش زمان آموزش همراه است. بنابراین، انتخاب اندازه مناسب مجموعه داده مستلزم ایجاد توازن میان دقت بازسازی و هزینه محاسباتی است.

جدول ۱. تأثیر اندازه مجموعه داده آموزشی بر عملکرد شبکه عصبی

| داده آموزشی | ۱۰۰۰ | ۲۰۰۰ | ۳۰۰۰ |
|-------------|------|------|------|
| معیار       |      |      |      |

۱۳. Pedrotti, F.L., L.M. Pedrotti, and L.S. Pedrotti, Introduction to optics. ۲۰۱۷: Cambridge University Press.
۱۴. Aime, C., E. Aristidi, and Y. Rabbia, The Fresnel diffraction: A story of light and darkness. European Astronomical Society Publications Series, ۲۰۱۳. ۵۹: p. ۳۷-۵۸.
۱۵. Sara, U., M. Akter, and M.S. Uddin, Image quality assessment through FSIM, SSIM, MSE and PSNR—a comparative study. Journal of Computer and Communications, ۲۰۱۹. ۷(۳): p. ۸-۱۸.
۱۶. Pinaya, W.H.L., et al., Autoencoders, in Machine learning. ۲۰۲۰, Elsevier. p. ۱۹۳-۲۰۸.
۱۷. Varotto, L., et al., Visual sensor network stimulation model identification via Gaussian mixture model and deep embedded features. Engineering Applications of Artificial Intelligence, ۲۰۲۲. ۱۱۴: p. ۱۰۵۰۹۶.
۱۸. Zhang, K., et al., *Deep image deblurring: A survey*. International Journal of Computer Vision, ۲۰۲۲. ۱۳۰(۹): p. ۲۱۰۳-۲۱۳۰.
۱۹. Zhang, Y. A better autoencoder for image: Convolutional autoencoder. in ICONIP۱۷-DCEC. Available online: <http://users.cecs.anu.edu.au/Tom.Gedeon/conf/ABCs۲۰۱۸/paper/ABCs۲۰۱۸>
۱. Lee, S., et al., Autofocusing and edge detection schemes in cell volume measurements with quantitative phase microscopy. Optics express, ۲۰۰۹. ۱۷(۸): p. ۶۴۷۶-۶۴۸۶.
۲. Kim, T., L. Yong, and Q. Xu. Robust design study on the wide angle lens with free distortion for mobile lens. in AOPC ۲۰۱۷: Optical Storage and Display Technology. ۲۰۱۷. SPIE.
۳. Petersen, B. Optical Anomalies and Lens Corrections Explained. ۲۰۱۶; Available from: <https://www.bhphotovideo.com/explora/photography/tips-and-solutions/optical-anomalies-and-lens-corrections-explained>.
۴. Islam, J., Towards AI-assisted disease diagnosis: learning deep feature representations for medical image analysis. Ph. D. dissertation, ۲۰۱۹.
۵. Zamir, S. W., et al. Restormer: Efficient Transformer for High-Resolution Image Restoration. CVPR, ۲۰۲۲, ۲۰۲۲.
۶. Liang, J., et al. SwinIR: Image Restoration Using Swin Transformer. ICCV Workshops, ۲۰۲۱.
۷. Mao, X., Shen, C., Yang, Y.-B. Image Restoration Using Very Deep Convolutional Encoder-Decoder Networks with Symmetric Skip Connections. NeurIPS, ۲۹, ۲۰۱۶.
۸. Ronneberger, O., Fischer, P., Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. MICCAI, ۲۳۴, ۲۰۱۵.
۹. Li, C., et al., Deep learning-based autofocus method enhances image quality in light-sheet fluorescence microscopy. Biomedical Optics Express, ۲۰۲۱. ۱۲(۸): p. ۵۲۱۴-۵۲۲۶.
۱۰. Sun, Y. and G. Mogos, Impact of Visual Distortion on Medical Images. IAENG International Journal of Computer Science, ۲۰۲۲. ۴۹(۱).
۱۱. Sun, Y., Predict the impact of visual distortion on medical images. BSc Thesis, Department of Computer Science and Software Engineering, Xi'an Jiaotong-Liverpool University, Suzhou, China, ۲۰۲۰.
۱۲. Wang, C., et al., Intelligent autofocus. arXiv preprint arXiv:۲۰۰۲.۱۲۳۸۹, ۲۰۲۰.

## Improving the quality of images with focus distortion using convolutional neural networks

Fatemeh Fallahi, Zahra Fallahi, Mohammadreza Jafarfard \*, Babak Zare

Physics Department, Iran University of Science and Technology, Iran

### Article details

Received: ۲۰۲۶/۰۲/۲۶

Accepted: ۲۰۲۶/۰۹/۱۰

Published: ۲۰۲۶/۰۹/۳۰

eISSN: ۲۵۸۸-۴۹۳x

ISSN: ۲۵۸۸-۴۸۲۱

Correspondence email:

[mrjafarfard@iust.ac.ir](mailto:mrjafarfard@iust.ac.ir)

DOI:

۱۰.۲۲۰۳۴/jmrph.۲۰۲۶.۱۲۲۹۲۳

### Abstract

Defocus distortion is one of the primary factors that degrades the quality of microscopic images, leading to reduced sharpness, loss of fine details, and decreased accuracy in image analysis. Therefore, the development of effective methods for defocus correction is of considerable importance in digital microscopy. In this study, a convolutional autoencoder neural network, consisting of convolutional layers, a compressed bottleneck, and transposed convolutional layers, was designed and trained to correct defocus distortion in digital microscopic images. To this end, a dataset comprising reference images and corresponding defocused images generated using the Fresnel diffraction model was constructed and employed for training and evaluation. The performance of the proposed model was assessed using the Mean Squared Error (MSE), Peak Signal-to-Noise Ratio (PSNR), and Structural Similarity Index (SSIM). The proposed model achieved an MSE of ۰.۰۰۰۸, a PSNR of ۳۰.۹۷ dB, and an SSIM of ۰.۹۸۳۹, indicating accurate image reconstruction and excellent structural similarity between the restored and reference images. These results demonstrate that the proposed approach provides an effective solution for defocus correction, improving the quality and sharpness of digital microscopic images.

**Keywords:** Focus, Fresnel diffraction, deep learning, convolutional neural network.