

## شناسایی شماره پلاک خودرو بر اساس یادگیری عمیق با نظارت ضعیف

محمد رضا محمدی\*

### چکیده

یادگیری عمیق در صورت استفاده از مجموعه داده آموزشی متناسب می‌تواند دقت قابل توجهی را نتیجه دهد. با این وجود، آماده‌سازی مجموعه داده بزرگ با حاشیه‌نویسی دقیق فرآیندی پرهزینه است. به همین دلیل، الگوریتم‌های یادگیری با نظارت ضعیف در سال‌های اخیر مورد توجه بسیاری از پژوهشگران قرار گرفته‌اند. در رویکرد یادگیری با نظارت ضعیف، مجموعه داده بزرگ اما با حاشیه‌نویسی ساده جمع‌آوری می‌شود تا هم نیازمندی شبکه‌های عمیق به داده‌های زیاد برآورده شود و هم هزینه حاشیه‌نویسی زیاد نباشد. در این مقاله، یک الگوریتم یادگیری با نظارت ضعیف برای شناسایی شماره پلاک خودرو با استفاده از شبکه‌های همگشتی عمیق پیشنهاد می‌شود. در فاز آموزش، تنها کاراکترهایی که در تصاویر پلاک وجود دارند مشخص می‌شود و الگوریتم پیشنهادی قادر است علاوه بر شناسایی وجود هر کاراکتر، مختصات آنها را نیز آشکارسازی نماید. بنابراین، در فاز آزمون شماره پلاک به صورت کامل قابل شناسایی است. به منظور ارزیابی عملکرد الگوریتم پیشنهادی، یک پایگاه داده شامل ۱۳۹۷ تصویر پلاک ایرانی جمع‌آوری شده است. نتایج آزمایش‌ها نشان می‌دهد ۹۵٫۶٪ از پلاک‌ها و ۹۹٫۱٪ از کاراکترها به درستی شناسایی شده‌اند.

### کلیدواژه‌ها

شماره پلاک خودرو، یادگیری عمیق، شبکه‌های همگشتی عمیق، یادگیری با نظارت ضعیف

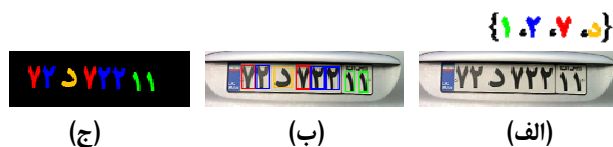
### ۱ مقدمه

بینایی ماشین از زمان شروع تصویربرداری رقمی در حوزه‌های مختلف صنعتی، کشاورزی و نظامی مورد استفاده قرار گرفته است. نخستین الگوریتم‌های حوزه بینایی ماشین به طور کامل مبتنی بر طراحی توسط یک مهندس خبره بوده‌اند. به عنوان نمونه، با استفاده از روش‌های پردازش تصویر میزان پر بودن یک بطری نوشیدنی در خط تولید قابل اندازه‌گیری است و می‌توان بر اساس

آن تصمیم‌گیری نمود. در این دوره الگوریتم‌های قدرتمندی مانند آستانه‌گذاری Otsu [۱] برای ناحیه‌بندی تصویر، تبدیل Hough [۲] برای تشخیص خط، و لبه‌یاب Canny [۳] معرفی شدند. با پیشرفت فن‌آوری تصویربرداری و افزایش دسترسی به آن، کاربردهای جدیدی با استفاده از بینایی کامپیوتر به وجود آمد. به عنوان نمونه، بازشناسی اعداد دستنویس (OCR) مورد توجه قرار گرفت. البته روش‌های ذکر شده برای جداسازی الگوهای اعداد پاسخگو نبودند. به همین دلیل، در این دوره الگوریتم‌های متنوعی مانند الگوهای دودویی محلی (LBP) [۴] و هیستوگرام جهت‌گرادیان (HOG) [۵] توسعه یافتند تا بتوانند محتوای تصویر را به خوبی توصیف کنند. سپس، با استفاده از الگوریتم‌های یادگیری ماشین مانند ماشین بردار پشتیبان (SVM) [۶] عملیات بازشناسی انجام می‌شد. تفاوت مهم میان روش‌های این دوره با دوره قبل در نیاز به داده‌های آموزشی است. در الگوریتم‌های دوره اول فرآیند

این مقاله در خردادماه ۱۳۹۸ دریافت، در بهمن‌ماه بازنگری و در اسفندماه همان سال پذیرفته شد.

\* دانشکده مهندسی کامپیوتر، دانشگاه علم و صنعت ایران  
رایانامه: mrmohammadi@iust.ac.ir



شکل ۱: روش‌های مرسوم حاشیه‌نویسی تصویر در کاربرد تشخیص اشیاء. الف) حاشیه‌نویسی در سطح تصویر، ب) حاشیه‌نویسی در سطح مستطیل محیطی، ج) حاشیه‌نویسی در سطح پیکسل

شکل ۱: روش‌های مرسوم حاشیه‌نویسی تصویر در کاربرد تشخیص اشیاء. الف) حاشیه‌نویسی در سطح تصویر، ب) حاشیه‌نویسی در سطح مستطیل محیطی، ج) حاشیه‌نویسی در سطح پیکسل [۱۵] ذکر شده است که حاشیه‌نویسی در سطح تصویر به طور میانگین ۱ ثانیه برای هر کلاس زمان نیاز دارد در حالیکه حاشیه‌نویسی در سطح مستطیل محیطی و پیکسل به ترتیب به ۱۰ و ۷۸ ثانیه زمان برای هر نمونه نیاز دارند. بنابراین، با صرف زمان یکسان، با استفاده از حاشیه‌نویسی سطح تصویر می‌توان مجموعه داده بسیار بزرگتری را جمع‌آوری نمود که برای آموزش مناسب شبکه‌های عمیق مورد نیاز است.

در این مقاله یک الگوریتم یادگیری با نظارت ضعیف برای شناسایی شماره پلاک خودروها پیشنهاد می‌شود. در گام آموزش تنها کاراکترهایی که در پلاک خودرو وجود دارند مشخص می‌شوند. الگوریتم پیشنهادی که مبتنی بر شبکه‌های کاملاً همگشتی<sup>۵</sup> است، می‌تواند بخشی از تصویر که متناسب با هر کاراکتر است را بیاموزد. در فاز آزمون، تصویر پلاک وارد می‌شود و موقعیت هر کدام از کاراکترهای آن مشخص می‌شود. بنابراین، برای تصاویر آزمون شماره پلاک به صورت کاملاً قابل شناسایی است.

دستاوردهای اصلی مقاله عبارتند از:

- تهیه یک پایگاه داده از تصاویر پلاک خودروهای ایرانی که شماره پلاک آنها در سطح تصویر حاشیه‌نویسی شده‌اند
- توسعه یک معماری مبتنی بر شبکه‌های کاملاً همگشتی عمیق برای شناسایی شماره پلاک خودرو با استفاده از تصاویر حاشیه‌نویسی شده در سطح تصویر (یادگیری با نظارت ضعیف)

ادامه مقاله به صورت زیر تدوین شده است. در بخش ۲ برخی از روش‌های موجود برای شناسایی شماره پلاک خودرو و آشکارسازی اشیاء مرور می‌شوند. روش پیشنهادی برای شناسایی شماره پلاک خودرو با استفاده از یادگیری با نظارت ضعیف در بخش ۳ ارائه می‌شود. پایگاه داده جمع‌آوری شده و نتایج استفاده از الگوریتم پیشنهادی در بخش ۴ گزارش می‌شود و جمع‌بندی مقاله در بخش ۵ انجام می‌شود.

## ۲ مرور کارهای پیشین

به دلیل کاربردهای فراوان، تشخیص و شناسایی شماره پلاک خودرو از سال‌ها پیش مورد توجه پژوهشگران در کشورهای

طراحی الگوریتم بینایی ماشین به طور کامل توسط مهندس خبره انجام می‌شد اما در روش‌های دوره دوم، تصمیم‌گیری نهایی توسط ماشین و با استفاده از داده‌های آموزشی انجام می‌شود.

ویژگی‌های مهندسی شده هر چند برای کاربردهای ساده بینایی ماشین مناسب بودند اما برای حل مسائل شناختی پیچیده عملکرد مطلوب را نداشتند. به همین دلیل، در سال‌های اخیر، الگوریتم‌های مبتنی بر یادگیری عمیق توسعه یافته‌اند که گام استخراج ویژگی نیز بر اساس داده‌های آموزشی و به صورت خودکار آموزش می‌بیند. این رویکرد از دهه ۱۹۹۰ با معرفی شبکه LeNet [۷] آغاز شد. با این وجود، به دلیل نیاز به داده‌های برجسب خورده زیاد و سخت‌افزارهای متناسب، چندان مورد توجه قرار نگرفت. در سال ۲۰۱۲، شبکه AlexNet [۸] توانست با اختلاف برنده چالش ILSVRC2012 [۹] شود. در سال‌های بعد، شبکه‌هایی عمیق‌تر و با کارایی بالاتر از جمله VGG [۱۰]، Inception [۱۱]، ResNet [۱۲] و DenseNet [۱۳] معرفی شدند و توانستند کارایی بالایی را در حوزه بینایی کامپیوتر نتیجه دهند.

با پیشرفت دانش بینایی ماشین، نتایج بدست آمده بهبود بسیاری یافته‌اند اما نیازمندی به داده‌های آموزشی نیز روز به روز بیشتر شده است. نیاز به داده‌های برجسب خورده زیاد در سال‌های اخیر استفاده از الگوریتم‌های یادگیری عمیق را با چالش روبرو کرده است. روش‌های انتقال یادگیری و داده‌افزایی در شبکه‌های عمیق باعث شده‌اند تا حدی مشکل کمبود داده برجسب خورده کاهش بیابد. با این وجود، توسعه الگوریتم‌هایی که بتوانند کارایی بالا را با استفاده از داده‌هایی که فرآیند برجسب‌گذاری آنها ساده باشد ادامه دارد. در حوزه یادگیری ماشین، به چنین الگوریتم‌هایی یادگیری با نظارت ضعیف<sup>۱</sup> گفته می‌شود.

الگوریتم‌های یادگیری با نظارت ضعیف به سه گروه اصلی تقسیم می‌شوند [۱۴]: نظارت غیرکامل<sup>۲</sup> که در آن برجسب بخش عمده‌ای از داده‌ها مشخص نیست، نظارت غیردقیق<sup>۳</sup> که در آن برجسب برای مجموعه‌ای از داده‌ها به صورت درهم مشخص است، و نظارت غیرصحیح<sup>۴</sup> که در آن برخی از برجسب‌های موجود ممکن است نادرست باشند. در هر سه حالت، جمع‌آوری مجموعه داده بسیار ساده‌تر از حالت با نظارت کامل است.

در حوزه آشکارسازی اشیاء در تصویر، حاشیه‌نویسی در سطح تصویر بجای حاشیه‌نویسی در سطح پیکسل یا مستطیل محیطی زیرمجموعه‌ای از روش‌های حاشیه‌نویسی غیردقیق به حساب می‌آید. این سه نوع حاشیه‌نویسی در شکل ۱ نشان داده شده است. در حالت حاشیه‌نویسی سطح تصویر، تنها اشیائی که در تصویر وجود دارند مشخص می‌شوند اما مکان آنها مشخص نمی‌شود و به همین دلیل بسیار سریع‌تر از حاشیه‌نویسی‌های دیگر است. در

<sup>1</sup> Weakly Supervised Learning

<sup>2</sup> Incomplete Supervision

<sup>3</sup> Inexact Supervision

<sup>4</sup> Inaccurate Supervision

<sup>5</sup> Fully Convolutional Networks

محاسبه امتیاز بیشینه، از یک تکنیک جایگزین با لحاظ کردن همسایگی مکانی استفاده شود.

با توجه به مزایای روش‌های یادگیری با نظارت ضعیف، در این مقاله یک شبکه عصبی همگشتی برای شناسایی شماره پلاک خودرو با استفاده از داده‌های حاشیه‌نویسی شده در سطح تصویر پیشنهاد می‌شود.

### ۳ روش پیشنهادی

در این مقاله یک روش مبتنی بر پنجره لغزان برای شناسایی کاراکترهای پلاک خودرو پیشنهاد شده است که به صورت نمادین در شکل ۲ نشان داده شده است. در گام پیش‌پردازش، با استفاده از روش متعادل‌سازی هیستوگرام وفقی کنتراست محدود<sup>۱</sup> (CLAHE) [۲۸] اثر نورپردازی تا حدی جبران می‌شود. از آنجایی که تصاویر پلاک می‌توانند در شرایط نوری بسیار متفاوتی ثبت شوند (به خصوص پلاک‌هایی که در شب تصویربرداری می‌شوند دارای کیفیت متفاوتی هستند)، کاهش اثر نورپردازی تاثیر زیادی در گام‌های بعد دارد.

پس از پیش‌پردازش، یک شبکه کاملاً همگشتی (FCN) برای شناسایی کاراکترهای موجود در تصویر ورودی طراحی شده است. در معماری پیشنهادی از ۹ لایه همگشتی با فیلترهای دارای ابعاد  $3 \times 3$  و ۱ لایه تلفیق حداکثر<sup>۲</sup> استفاده شده است که جزئیات آن در جدول ۱ گزارش شده است. با الگوبرداری از معماری VGG، در شبکه طراحی شده تنها از فیلترهای با ابعاد  $3 \times 3$  استفاده شده است. همچنین، تعداد لایه‌ها به گونه‌ای انتخاب شده است که میدان تاثیر<sup>۳</sup> لایه آخر دارای ابعاد  $33 \times 33$  باشد (بزرگترین طول و عرض کاراکترها در پلاکی با ابعاد  $50 \times 250$  پیکسل حدود  $33$  پیکسل است).

در دو لایه همگشتی نخست، ویژگی‌های سطح پائینی از تصویر استخراج می‌شوند. سپس، با استفاده از تلفیق حداکثر، این ویژگی‌ها در بعد مکان کاهش می‌یابند. سپس، توسط ۶ لایه همگشتی، این ویژگی‌ها ترکیب می‌شوند تا ویژگی‌های سطح بالاتری محاسبه شوند. در نهایت توسط یک لایه همگشتی که دارای ۲۷ فیلتر (به تعداد کاراکترها) با تابع فعال‌سازی Sigmoid است، عمل طبقه‌بندی انجام می‌شود. ورودی این شبکه یک تصویر رنگی با ابعاد  $50 \times 250$  است و خروجی ۲۷ نقشه فعالیت با ابعاد  $9 \times 109$  است. در لایه‌های همگشتی این شبکه از zero-padding استفاده نشده است زیرا صفرهای غیرواقعی به تصویر اضافه می‌کند و علاوه بر افزایش حجم محاسبات، اثر مخربی بر آموزش شبکه دارد. به همین دلیل، پس از هر لایه همگشتی  $3 \times 3$ ، ابعاد نقشه‌های فعالیت ۲ واحد کاهش می‌یابد. در ادامه، فاز آموزش و آزمون شبکه پیشنهادی بررسی می‌شود.

مختلف قرار گرفته است. شناسایی شماره پلاک خودرو به طور معمول شامل دو گام ناحیه‌بندی کاراکترها و طبقه‌بندی آنها است. هدف از گام ناحیه‌بندی، یافتن محدوده مربوط به هر کدام از کاراکترهای پلاک خودرو است. با توجه به متمایز بودن رنگ کاراکترهای پلاک از پس‌زمینه پلاک، در برخی مطالعات مانند [۱۶] از روش‌های ساده آستانه‌گذاری و عملگرهای مورفولوژی برای ناحیه‌بندی تصویر پلاک استفاده شده است. با وجود عملکرد مناسب چنین الگوریتم‌هایی در تصاویر ساده، شرایطی مانند انسداد جزئی، محوشدگی، وجود سایه و کثیف بودن تاثیر منفی زیادی بر عملکرد آنها دارد. در [۱۷، ۱۸] از هیستوگرام‌های افقی و عمودی برای مشخص کردن دقیق‌تر مستطیل‌های محیطی کاراکترهای پلاک استفاده شده است. با پیشرفت‌های بدست آمده توسط یادگیری عمیق، از این روش برای ناحیه‌بندی دقیق‌تر کاراکترهای پلاک در [۱۹] استفاده شده است.

پس از مشخص کردن ناحیه مربوط به هر کدام از کاراکترهای پلاک، نیاز است تا کاراکتر مربوطه مشخص شود. در [۲۰، ۲۱] از انطباق ویژگی‌های SIFT برای بازشناسی کاراکترهای پلاک استفاده شده است. در [۲۲] ویژگی‌های ساده‌ای مانند میانگین محلی پیکسل‌ها از هر کاراکتر استخراج شده و طبقه‌بندی توسط شبکه عصبی انجام شده است. در بسیاری از کارهای اخیر از جمله [۲۳، ۲۴] از شبکه‌های عصبی همگشتی برای استخراج ویژگی و طبقه‌بندی کاراکترها استفاده شده است.

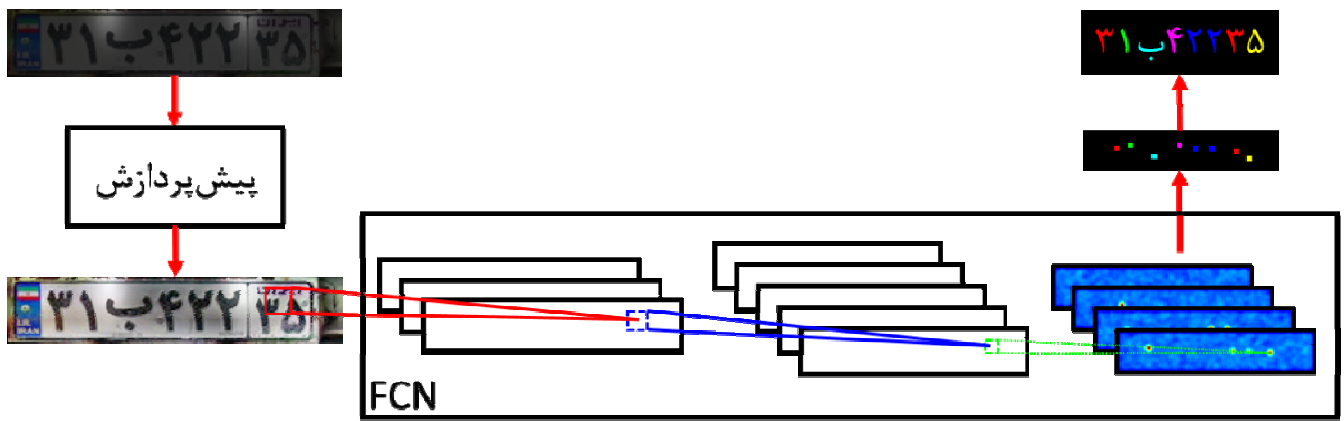
یادگیری عمیق، به ویژه شبکه‌های عصبی همگشتی، در قسمت‌های مختلف از فرآیند بازشناسی شماره پلاک خودرو به طور موفقیت‌آمیزی استفاده شده است. با این وجود، نیاز به حجم زیادی از تصاویر حاشیه‌نویسی شده استفاده از آن را با چالش روبرو کرده است. به خصوص باید در نظر داشت پلاک خودرو در کشورهای مختلف دارای قالب‌ها و زبان‌های متفاوتی است، و نیاز است در هر کشور مجموعه داده مورد نظر به صورت مستقل جمع‌آوری شود. در سال‌های اخیر روش‌هایی با نام یادگیری با نظارت ضعیف مورد توجه بیشتری قرار گرفته‌اند تا بتوان هزینه حاشیه‌نویسی را کاهش داد. از آنجایی که الگوریتم‌های یادگیری عمیق نیازمند داده‌های زیادی هستند، با استفاده از حاشیه‌نویسی سطح تصویر می‌توان تعداد بسیار بیشتری از تصاویر را با صرف زمان یکسان حاشیه‌نویسی کرد.

در [۲۵، ۲۶] یک شبکه کاملاً همگشتی برای تشخیص اشیاء در تصویر پیشنهاد شده است. در این معماری، تصویر وارد تعدادی لایه همگشتی و تلفیق می‌شود و دسته مربوط به هر شی نیز توسط لایه‌های همگشتی تخمین زده می‌شود (بجای لایه‌های کاملاً متصل). به منظور توسعه شبکه پیشنهادی برای تشخیص اشیاء در ابعاد مختلف، پیشنهاد شده است تا تصویر با ابعاد مختلف به چندین شبکه موازی وارد شود. یکی از چالش‌های موجود در الگوریتم‌های تشخیص اشیاء با نظارت ضعیف، تعیین خودکار محدوده اشیاء است. در [۲۷] پیشنهاد شده است تا بجای

<sup>۱</sup> Contrast Limited Adaptive Histogram Equalization

<sup>۲</sup> Max-Pooling

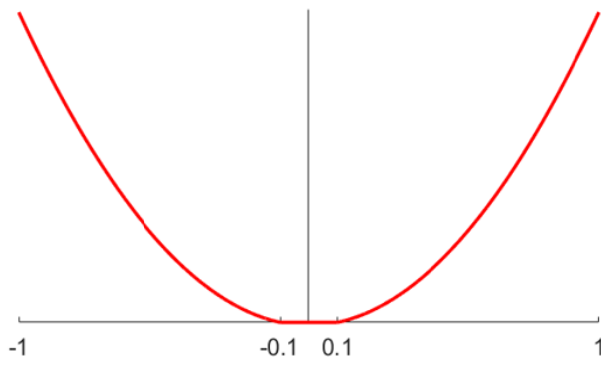
<sup>۳</sup> Receptive Field



شکل ۲: نمای شماتیک الگوریتم پیشنهادی برای شناسایی شماره پلاک خودرو

استفاده از ایده GlobalMaxPooling در فاز آموزش به این معنا است که برای هر کاراکتر، در هر تصویر تنها مختصاتی که دارای بالاترین امتیاز بوده است در تغییر وزن‌های شبکه دخالت داده می‌شود. به عبارت دیگر، برچسب سطح تصویر برای هر کاراکتر تنها برای مستطیلی در نظر گرفته می‌شود که بزرگترین پیش‌بینی را برای آن کاراکتر داشته است. سپس، بسته به وجود یا عدم وجود این کاراکتر در تصویر، وزن‌های شبکه به گونه‌ای اصلاح می‌شوند که پیش‌بینی برای این مستطیل تقویت یا تضعیف شود. در تکرار نخست مختصات دارای بیشترین امتیاز برای هر کاراکتر تصادفی خواهد بود اما به تدریج به ناحیه‌ای همگرا خواهد شد که در تصاویر مثبت وجود دارد اما در تصاویر منفی مشاهده نمی‌شود. در فاز آزمون، تصویر ورودی به شبکه وارد می‌شود و ۲۷ نقشه فعالیت محاسبه می‌شوند. خروجی شبکه برای هر کاراکتر یک ماتریس  $9 \times 109$  است که احتمال وجود کاراکتر مورد نظر در ۹۸۱ موقعیت را تخمین می‌زند. یک کاراکتر از پلاک ممکن است منجر به فعال شدن چندین مقدار نزدیک به هم شود. به همین دلیل، در هر نقشه ویژگی از عملیات حذف مقادیر غیربیشینه<sup>۱</sup> (NMS) با ابعاد  $9 \times 9$  استفاده می‌شود تا یک کاراکتر به اشتباه چند مرتبه تشخیص داده نشود. به عبارت دیگر، تنها مقادیری که در همسایگی  $9 \times 9$  خود دارای بیشترین مقدار هستند در ادامه بررسی می‌شوند.

$$loss = \max(0, (y_{true} - y_{pred})^2 - 0.1^2)$$



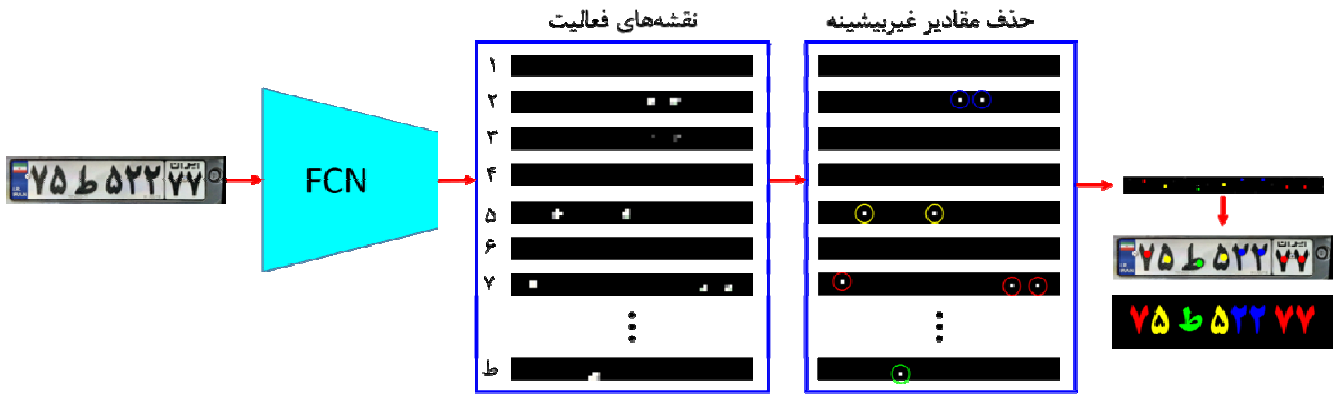
شکل ۳: تابع اتلاف مورد استفاده در فرآیند بهینه‌سازی

جدول ۱: ساختار شبکه کاملاً همگشتی پیشنهادی

نوع لایه	تابع فعال‌سازی	ابعاد خروجی	تعداد پارامتر
ورودی	-	$50 \times 250 \times 3$	۰
Conv2D	ReLU	$48 \times 248 \times 64$	۱۷۹۲
Conv2D	ReLU	$46 \times 246 \times 64$	۳۶۹۲۸
MaxPool	-	$23 \times 123 \times 64$	۰
Conv2D	ReLU	$21 \times 121 \times 128$	۷۳۸۵۶
Conv2D	ReLU	$19 \times 119 \times 128$	۱۴۷۵۸۴
Conv2D	ReLU	$17 \times 117 \times 256$	۲۹۵۱۶۸
Conv2D	ReLU	$15 \times 115 \times 256$	۵۹۰۰۸۰
Conv2D	ReLU	$13 \times 113 \times 256$	۵۹۰۰۸۰
Conv2D	ReLU	$11 \times 111 \times 256$	۵۹۰۰۸۰
Conv2D	Sigmoid	$9 \times 109 \times 27$	۶۲۲۳۵
مجموع			۲۳۸۷۸۰۳

در فاز آموزش نیاز است تا پیش‌بینی شبکه برای ۲۷ کاراکتر با حاشیه نوشته شده توسط انسان مقایسه شود. از آنجایی که حاشیه‌نویسی در سطح تصویر انجام شده است، تنها وجود یا عدم وجود یک کاراکتر در تصویر مشخص است. بنابراین، نیاز است تا نقشه فعالیت مربوط به هر کاراکتر از یک ماتریس  $9 \times 109$  به یک عدد تبدیل شود. برای هر کاراکتر، مختصاتی که دارای بالاترین امتیاز باشد به عنوان پیش‌بینی شبکه برای آن کاراکتر در نظر گرفته می‌شود. این عمل معادل با آن است که در فاز آموزش پس از لایه انتهایی یک لایه GlobalMaxPooling قرار داده شود تا بالاترین امتیاز برای هر کاراکتر محاسبه شود. پس از محاسبه پیش‌بینی شبکه برای هر کاراکتر، نیاز است تا خطای آن نسبت به مقدار واقعی (۱ برای کاراکتری که وجود دارد و ۰ برای کاراکتری که وجود ندارد) محاسبه شود. برای تابع اتلاف از مقدار مربع خطای برش خورده استفاده می‌شود (شکل ۳). در این تابع اتلاف،  $y_{true}$  و  $y_{pred}$  به ترتیب مقادیر مطلوب و پیش‌بینی شده برای یک کاراکتر هستند و از خطای کمتر از ۰٫۱ صرف نظر می‌شود.

<sup>۱</sup> Non-Maximum Suppression



شکل ۴: فرآیند شناسایی شماره پلاک خودرو برای تصاویر آزمون

می‌شود. در فاز آموزش، تنها وجود یا عدم وجود هر کاراکتر در اختیار شبکه قرار می‌گیرد اما برای ارزیابی عملکرد شبکه در مکان‌یابی کاراکترها، شماره پلاک خوانده شده با شماره‌ای که توسط خبره ثبت شده است مقایسه می‌شود. به عبارت دیگر، آموزش شبکه به صورت یادگیری با نظارت ضعیف انجام می‌شود اما برای ارزیابی از نظارت کامل استفاده می‌شود.

هیستوگرام مربوط به ۲۷ کاراکتری که در این مجموعه داده مشاهده شده است در شکل ۶ ترسیم شده است. همانطور که مشاهده می‌شود، کاراکترهای ۱ تا ۹ دارای بیشترین تکرار در پایگاه داده هستند. پلاک‌های استاندارد ایران دارای ۸ کاراکتر هستند که ۷ مورد از آنها عدد و یک مورد حرف است. همچنین، در حال حاضر عدد ۰ تنها در آخرین جایگاه می‌تواند وجود داشته باشد و به همین دلیل، مشاهده زیاد کاراکترهای ۱ تا ۹ طبیعی است. در نقطه مقابل کاراکترهای الف (خودروهای دولتی)، □ (خودروهای معلولین) و ۰ به ترتیب با ۱۴، ۱۸ و ۳۱ کمترین تکرار را داشته‌اند.

#### ۲-۴ معیارهای ارزیابی

مقدار اتلاف نهایی حاصل از الگوریتم پیشنهادی نشان دهنده میزان توانایی شبکه در یادگیری تشخیص وجود هر کاراکتر در تصویر است. با این وجود، در یادگیری با نظارت ضعیف به دنبال این هستیم که الگوریتم پیشنهادی بتواند مکان کاراکترها در تصویر را نیز محاسبه نماید. بنابراین، شماره شناسایی شده برای هر پلاک به طور کامل نسبت به شماره صحیح آن سنجیده می‌شود و درصد پلاک‌هایی که تمام کاراکترهای آن به درستی شناسایی شده‌اند به عنوان معیار  $Acc_{plate}$  گزارش می‌شود. از آنجایی که در هر تصویر ۸ کاراکتر وجود دارد و اشتباه شدن هر کدام منجر به اشتباه بودن شماره پلاک می‌شود، علاوه بر معیار فوق، دقت در سطح کاراکتر ( $Acc_{char}$ ) نیز گزارش خواهد شد. برای محاسبه  $Acc_{char}$ ، ابتدا فاصله Levenshtein [۲۹] بین پلاک تشخیص داده شده و مقدار صحیح آن محاسبه می‌شود تا تعداد کاراکترهای اشتباه بدست بیاید. سپس، مجموع کاراکترهای اشتباه به کل کاراکترها محاسبه شده و اختلاف آن با ۱ به عنوان  $Acc_{char}$  اعلام می‌شود.

علاوه بر این، ممکن است یک کاراکتر در چند نقشه ویژگی متفاوت فعال شود (به عنوان نمونه، کاراکتر "۲" به کاراکتر "۳" شباهت دارد و امکان دارد خروجی شبکه در یک ناحیه مربوط به کاراکتر "۲"، دارای احتمال بالا برای کاراکتر "۳" نیز باشد). به همین دلیل، عملیات حذف مقادیر غیربیشینه بین نقشه‌های ویژگی نیز با ابعاد  $9 \times 5$  انجام می‌شود. به عبارت دیگر، تنها مقادیری که در همسایگی  $9 \times 5 \times 27$  خود دارای بیشترین مقدار هستند در ادامه بررسی می‌شوند.

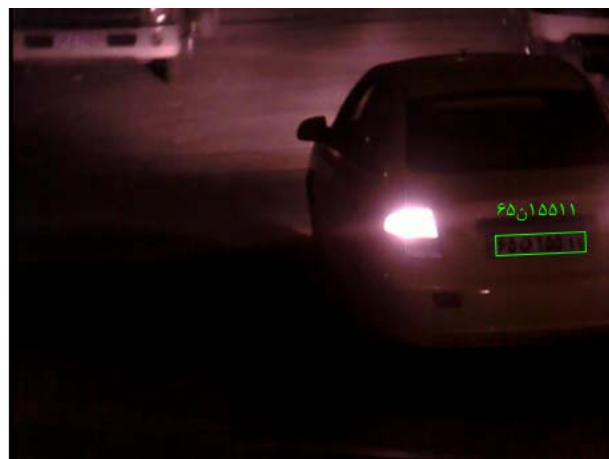
پس از حذف مقادیر غیربیشینه، مقادیری که از یک حد آستانه بزرگتر باشند به عنوان وجود کاراکتر معرفی می‌شوند. در نهایت، کاراکترهای تشخیص داده شده بر اساس مختصات افقی مرتب شده و شماره پلاک گزارش می‌شود. با استفاده از این استراتژی یادگیری با نظارت ضعیف، در فاز آموزش تنها برچسب در سطح تصویر وجود دارد، اما در فاز آزمون مختصات کاراکترها و شماره پلاک نیز مشخص می‌شوند. این عملیات به طور نمادین در شکل ۴ به تصویر کشیده شده است. در این شکل، خروجی FCN پیشنهادی ۲۷ نقشه ویژگی است. پس از حذف مقادیر غیربیشینه، نقاط دارای احتمال بالا با دایره مشخص شده‌اند که از کنار هم قرار دادن کاراکترهای مربوطه، شماره پلاک خوانده می‌شود.

#### ۴ نتایج آزمایش‌ها

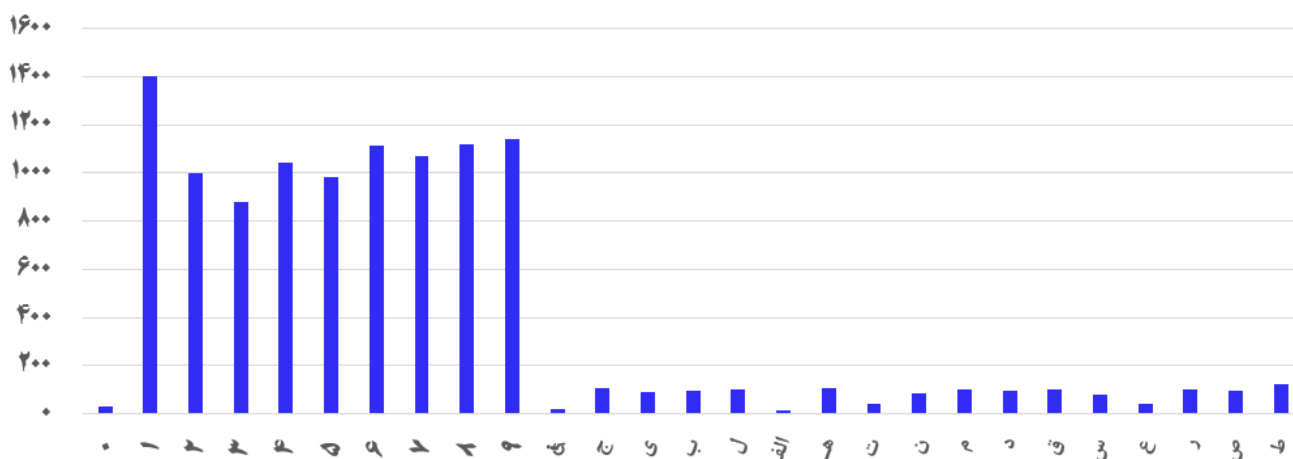
در این بخش ابتدا پایگاه داده مورد استفاده برای ارزیابی الگوریتم پیشنهادی تشریح می‌شود. سپس، جزئیات پیاده‌سازی ذکر می‌شود. در ادامه معیارهای مورد استفاده برای ارزیابی توضیح داده شده و نتایج بدست آمده گزارش و تحلیل می‌شوند.

#### ۴-۱ پایگاه داده

به منظور ارزیابی الگوریتم پیشنهادی، یک پایگاه داده شامل ۱۳۴۰ تصویر و ۱۳۹۷ پلاک از دوربین‌های مختلف و در شرایط آب و هوایی و نورپردازی‌های متنوع جمع‌آوری شده است. برای هر پلاک، مختصات چهار گوشه و شماره پلاک توسط یک خبره حاشیه‌نویسی شده است که نمونه‌هایی از آنها در شکل ۵ نشان داده شده است. هدف از این پژوهش شناسایی شماره پلاک خودرو بوده است و بنابراین، مختصات چهار گوشه پلاک دانسته فرض



شکل ۵: نمونه‌هایی از داده‌های مورد استفاده برای شناسایی شماره پلاک خودرو



شکل ۶: تعداد هر کدام از کاراکترهای مورد بررسی در پایگاه داده جمع‌آوری شده

که  $x_k$  و  $y_k$  به ترتیب مختصات افقی و عمودی  $k$  امین گوشه پلاک هستند،  $width$  و  $height$  به ترتیب عرض و ارتفاع پلاک هستند، و  $rand$  یک متغیر تصادفی با توزیع مثالی بین ۳ نقطه است. به عبارت دیگر، هر کدام از چهار گوشه حداکثر به اندازه ۱۰ درصد از ابعاد پلاک جابجا می‌شوند. سپس، این چهار گوشه توسط تبدیل  $perspective$  به چهار نقطه ثابت نگاشت می‌شوند تا یک تصویر  $۲۵۰ \times ۵۰$  پیکسل بدست بیاید.

از آنجایی که کاراکترهای پلاک در خودروهای ایرانی می‌توانند روشن یا تیره باشند و همچنین رنگ زمینه پلاک متنوع است (زرد، سفید و قرمز)، در هر تکرار به صورت تصادفی ترتیب ۳ کانال رنگی تصویر جابجا شده و هر کانال با احتمال ۵۰٪ قرینه می‌شود. اصولاً برای تشخیص کاراکتر هندسه آن اهمیت دارد و رنگ کاراکتر کم‌اهمیت است. علاوه بر این، به هر تصویر به صورت تصادفی مقداری نویز و بایاس اضافه می‌شود و همچنین، یک فیلتر گاوسی با انحراف معیار تصادفی بر روی تصویر اعمال می‌شود. با استفاده از این عملیات، داده‌افزایی مناسبی انجام می‌شود که کیفیت آموزش شبکه را بهبود می‌دهد و از بیش‌برازش آن جلوگیری می‌کند.

#### ۳-۴ جزئیات پیاده‌سازی

با توجه به اینکه حجم پایگاه داده به نسبت کوچک است و به خصوص تعداد مشاهده برخی کاراکترها بسیار کم است، برای ارزیابی الگوریتم پیشنهادی از روش ارزیابی متقابل با ۵ فولد استفاده می‌کنیم. در این روش ارزیابی، تصاویر پلاک‌ها به ۵ گروه تقسیم می‌شوند. سپس، در هر آزمایش از تصاویر ۴ گروه برای آموزش شبکه و از تصاویر گروه دیگر برای ارزیابی شبکه استفاده می‌شود. پس از انجام ۵ آزمایش، نتایج ارزیابی آنها با همدیگر تلفیق می‌شوند تا عملکرد الگوریتم پیشنهادی به درستی سنجیده شود.

یک راه‌کار مناسب برای بهبود کیفیت آموزش شبکه با داده‌های کم استفاده از داده‌افزایی<sup>۱</sup> است. برای داده‌افزایی هندسی، مختصات چهارگوشه هر پلاک به صورت تصادفی و طبق رابطه زیر جابجا می‌شوند:

$$x_k = x_k + rand(-0.1, 0, 0.1) \times width \quad (1)$$

$$y_k = y_k + rand(-0.1, 0, 0.1) \times height \quad (2)$$

<sup>1</sup> Data Augmentation

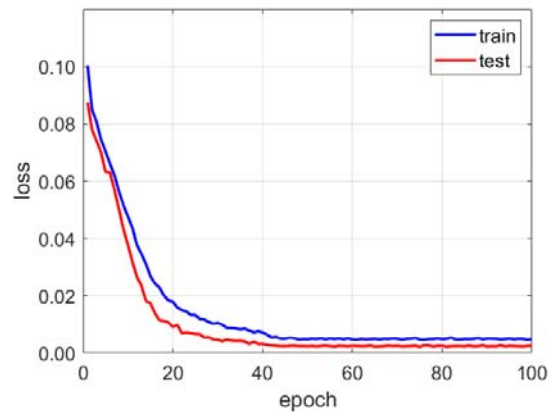
کاراکتر "۷" در تصویر پلاک است. مقدار بیشترین فعالیت برای ۲۷ نقشه در ستون آخر از این شکل نشان داده شده است که دارای هشت نقطه روشن مربوط به هشت کاراکتر پلاک است. در این شکل، نقاط متناظر بر روی تصویر پلاک نیز ترسیم شده است و کاراکترهای تشخیص داده شده که بر اساس مختصات افقی مرتب شده‌اند نیز نوشته شده است. یک نکته با اهمیت در این شکل آن است که مکان کاراکترها در فاز آموزش استفاده نشده است، اما الگوریتم پیشنهادی به خوبی توانسته است ویژگی متمایز برای هر کاراکتر را شناسایی نماید.

دقت الگوریتم پیشنهادی در سطح شناسایی صحیح پلاک و کاراکترهای آن در جدول ۲ گزارش شده است. همانطور که مشاهده می‌شود، در مجموع شماره ۱۳۳۶ پلاک از ۱۳۹۷ مورد به طور کامل به درستی شناسایی شده‌اند که معادل با دقت ۹۵,۶ درصد است. از میان ۶۱ پلاکی که شماره آنها با خطا شناسایی شده است، تنها ۱۰۱ کاراکتر اشتباه شده‌اند که نسبت به ۱۱۱۷۶ کاراکتر معادل با دقت ۹۹,۱ درصد در سطح تشخیص کاراکتر است. از ۱۰۱ پلاک دارای خطا، ۳۳ پلاک دارای تنها ۱ کاراکتر اشتباه بوده‌اند. تعداد پلاک‌های دارای ۲، ۳، ۴ و ۵ خطا نیز به ترتیب ۲۰، ۵، ۲ و ۱ بوده است. هیچ پلاکی نیز دارای خطای بیش از ۵ کاراکتر نبوده است.

در شکل ۸ نمونه‌هایی از نتایج الگوریتم پیشنهادی برای شناسایی شماره پلاک خودرو نشان داده شده است. ستون اول در این شکل تصاویری را نشان می‌دهد که پس از پیش‌پردازش وارد بخش FCN می‌شوند. مشابه با شکل ۴، خروجی FCN برای هر تصویر ۲۷ نقشه فعالیت است. در ستون دوم از شکل ۸ ترکیب (ماکزیمم) این نقشه‌ها نشان داده شده است. مختصات معادل با نقاطی که دارای امتیاز بیش از حد آستانه هستند در ستون سوم با دایره بر روی تصویر اصلی نشان داده شده است. در ستون‌های چهارم و پنجم نیز به ترتیب شماره شناسایی شده و شماره صحیح نوشته شده است. در این دو ستون، کاراکترهایی که اشتباه شده‌اند با رنگ قرمز متمایز هستند. این نتایج نشان می‌دهند که الگوریتم پیشنهادی به خوبی توانسته است پلاک‌هایی با رنگ‌ها و کیفیت‌های متنوع را شناسایی کند.

جدول ۲: عملکرد الگوریتم پیشنهادی

کاراکتر	پلاک			کاراکتر		
	صحیح	اشتباه	دقت	صحیح	اشتباه	دقت
فولد ۱	۲۶۷	۱۳	۹۵,۴	۲۲۲۱	۱۹	۹۹,۲
فولد ۲	۲۶۱	۱۹	۹۳,۲	۲۲۰۷	۳۳	۹۸,۵
فولد ۳	۲۶۸	۱۱	۹۶,۱	۲۲۱۳	۱۹	۹۹,۱
فولد ۴	۲۷۴	۵	۹۸,۲	۲۲۲۴	۸	۹۹,۶
فولد ۵	۲۶۶	۱۳	۹۵,۳	۲۲۱۰	۲۲	۹۹,۰
مجموع	۱۳۳۶	۶۱	۹۵,۶	۱۱۰۷۵	۱۰۱	۹۹,۱



شکل ۷: نمودار تابع اتلاف شبکه پیشنهادی برای فولد نخست

برای پیاده‌سازی الگوریتم پیشنهادی از کتابخانه Keras استفاده شده است. با توجه به اهمیت لایه‌های ابتدایی در استخراج ویژگی‌های همه منظوره، وزن‌های چهار لایه همگشتی ابتدایی از شبکه VGG16 [۱۰] که بر روی مجموعه داده ImageNet [۳۰] آموزش دیده‌اند به شبکه پیشنهادی منتقل شده‌اند. سپس، تمام وزن‌های شبکه با نرخ یادگیری ۰,۰۰۰۱ در ۱۰۰ دوره<sup>۱</sup> توسط بهینه‌ساز Adam آموزش دیده‌اند.

با توجه به غیرمتوازن بودن تعداد تکرار کاراکترهای مختلف، در تابع اتلاف برای هر کاراکتر یک وزن در نظر گرفته شده است. وزن کاراکترهایی که در یک پلاک حضور ندارند برابر با  $w_k$  و برای کاراکترهایی که حضور دارند  $1 - w_k$  است که برابر با درصد پلاک‌هایی است که حاوی  $k$  امین کاراکتر هستند.

زمان اجرای شبکه برای هر پلاک بر روی یک سیستم دارای پردازنده Core i7-7700HQ CPU @ 2.80GHz به طور میانگین ۱۶ هزارم ثانیه بوده است. زمان اجرای الگوریتم NMS و خواندن پلاک نیز به طور میانگین ۱۹ هزارم ثانیه بوده است. بنابراین، در مجموع حدود ۳۵ هزارم ثانیه برای خواندن هر پلاک با استفاده از یک پردازنده معمولی زمان نیاز است که استفاده از آن در سیستم‌های بی‌درنگ را ممکن می‌سازد.

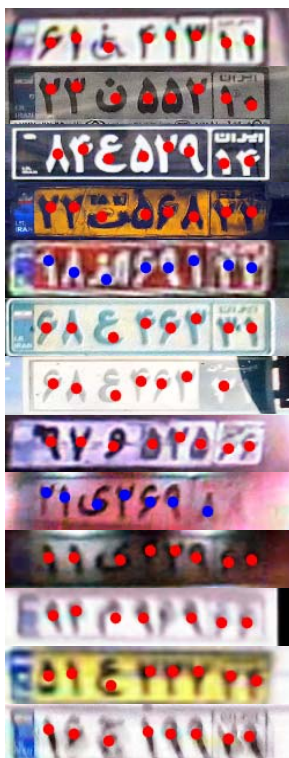
#### ۴-۴ نتایج

نمودار تابع اتلاف بر حسب دوره آموزش در شکل ۷ برای فولد نخست ترسیم شده است. همانطور که مشاهده می‌شود، مقدار تابع اتلاف برای هر دو بخش آموزش و آزمون به خوبی کاهش می‌یابد که نشان می‌دهد یادگیری شبکه به خوبی انجام می‌شود.

فرآیند شناسایی شماره پلاک خودرو برای تصاویر آزمون در شکل ۴ نشان داده شده است. تصویر ورودی پس از پیش‌پردازش وارد FCN آموزش دیده در فاز آموزش می‌شود و خروجی آن ۲۷ نقشه فعالیت با ابعاد  $9 \times 109$  است. در هر نقشه فعالیت، مقادیر نزدیک به یک و صفر به ترتیب معادل با وجود و عدم وجود کاراکتر در آن نقطه هستند. به عنوان نمونه، در نقشه فعالیت مربوط به کاراکتر "۷" سه ناحیه روشن دیده می‌شود که معادل با سه

<sup>۱</sup> epoch

۶۱ □ ۴۱۳۱۱	۶۱ □ ۴۱۳۱۱
۲۳ ن ۵۵۲۱۰	۲۳ ن ۵۵۲۱۰
۸۴ ع ۵۲۹۱۴	۸۴ ع ۵۲۹۱۴
۲۲ ت ۵۶۸۲۲	۲۲ ت ۵۶۸۲۲
۱۸ الف ۶۹۱۲۲	۱۸ الف ۶۹۱۲۲
۶۸ ع ۴۶۳۳۱	۶۸ ع ۴۶۳۳۱
۶۸ ع ۴۶۳۳۱	۶۸ ع ۴۶۳۳۱
۹۷ و ۵۲۵۶۶	۹۷ و ۵۲۵۶۶
۲۱ ی ۲۶۹۸۸	۲۱ ی ۲۶۹۸۸
۱۱ ی ۴۲۹۶۶	۱۱ ی ۴۲۹۶۶
۹۴ م ۹۶۹۶۶	۹۴ م ۹۶۹۶۶
۵۱ ع ۲۳۷۲۴	۵۱ ع ۲۳۷۲۴
۱۶ ج ۱۹۹۷۹	۱۶ ج ۱۹۹۷۹
شماره صحیح	شماره شناسایی شده



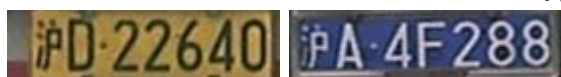
بزرگترین امتیازها کاراکترهای آشکار شده تصویر ورودی شبکه

شکل ۸: نمونه‌هایی از عملکرد الگوریتم پیشنهادی برای شناسایی شماره پلاک خودرو

#### ۴-۵ مقایسه با کارهای پیشین

به دلیل عدم دسترسی به یک مجموعه داده عمومی از پلاک‌های ایرانی، به منظور مقایسه با کارهای پیشین از مجموعه داده LPRNet [۳۱] استفاده می‌کنیم. این مجموعه داده شامل تصاویر بریده شده از پلاک‌های چینی است که دارای هفت کاراکتر (حرف چینی، حرف انگلیسی و رقم انگلیسی) است و نمونه‌هایی از آن در شکل ۱۰ نشان داده شده است.

با وجود آنکه در نسخه اصلی LPRNet تعداد ۱۱۶۹۶ تصویر وجود دارد، در نسخه در دسترس آن تنها ۷۵۲ تصویر آموزشی و ۱۵۰ تصویر ارزیابی وجود دارد. در این تصاویر، ۵۵ کاراکتر مختلف (۱۰ عدد، ۲۴ حرف انگلیسی و ۳۱ حرف چینی) وجود دارد. کاراکتر 苏 با ۴۳۹ داده آموزشی بیشترین و کاراکترهای 津، 蒙، 贵، 湘، 吉 و 辽 با تنها ۱ داده آموزشی کمترین تکرار را دارا هستند. نتایج حاصل از الگوریتم پیشنهادی در جدول ۳ گزارش شده است. همانطور که مشاهده می‌شود، الگوریتم پیشنهادی توانسته است ۹۲٪ از پلاک‌ها (۱۳۸ از ۱۵۰) را به درستی بخواند در حالیکه دقت الگوریتم LPRNet برابر با ۹۵٪ است. یکی از دلایل دقت بالاتر LPRNet آن است که با استفاده از داده‌های بسیار بیشتری آموزش دیده است. به طور خاص، آموختن کاراکترهایی با ۱ داده آموزشی به صورت با نظارت ضعیف بسیار دشوار است.



نکته قابل توجهی که در شکل ۸ به وضوح مشاهده می‌شود، یادگیری نقطه متمایزگر برای هر کدام از کاراکترها توسط شبکه پیشنهادی است. به عنوان نمونه، برای شناسایی کاراکتر "۵" نقطه‌ای نزدیک به مرکز، و برای کاراکتر "۳" نقطه‌ای نزدیک به دندانه‌ها توسط شبکه انتخاب شده است.

در الگوریتم پیشنهادی، از ایده پنجره لغزان برای شناسایی کاراکترها توسط FCN استفاده شده است. هر کدام از نقاط در نقشه فعالیت تنها پیکسل‌های درون یک پنجره  $33 \times 33$  را لحاظ می‌کنند. به منظور نمایش مشخصه‌هایی که توسط شبکه آموخته شده است، میانگین ناحیه‌هایی از تصاویر آزمون که برای هر کدام از ۲۷ کاراکتر منجر به امتیاز بیش از ۰٫۵ در خروجی شده است در شکل ۹ به تصویر کشیده شده است. همانطور که مشاهده می‌شود، الگوریتم پیشنهادی توانسته است به خوبی مشخصه‌های متمایزگر کاراکترها را بیاموزد.



شکل ۹: میانگین محدوده‌هایی از تصاویر که با امتیاز بالا به عنوان هر کاراکتر شناسایی شده‌اند



- [3] J. Canny, "A Computational Approach to Edge Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679–698, 1986.
- [4] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern Recognition*, vol. 29, no. 1, pp. 51–59, 1996.
- [5] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2005, pp. 886–893.
- [6] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
- [7] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [8] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.
- [9] O. Russakovsky et al., "Imagenet large scale visual recognition challenge," *International Journal of Computer Vision*, vol. 115, no. 3, pp. 211–252, 2015.
- [10] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
- [11] C. Szegedy et al., "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1–9.
- [12] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
- [13] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708.
- [14] Z.-H. Zhou, "A brief introduction to weakly supervised learning," *National Science Review*, vol. 5, no. 1, pp. 44–53, 2017.
- [15] A. Bearman, O. Russakovsky, V. Ferrari, and L. Fei-Fei, "What's the point: Semantic segmentation with point supervision," in *European conference on computer vision*, 2016, pp. 549–565.
- [16] M. S. Sarfraz, A. Shahzad, M. A. Elahi, M. Fraz, I. Zafar, and E. A. Edirisinghe, "Real-time automatic license plate recognition for CCTV forensic applications," *Journal of real-time image processing*, vol. 8, no. 3, pp. 285–295, 2013.
- [17] M. Nejati, A. Majidi, and M. Jalalat, "License plate recognition based on edge histogram analysis and



شکل ۱۰: چند تصویر نمونه از مجموعه داده LPRNet

جدول ۳: نتایج حاصل از الگوریتم پیشنهادی و الگوریتم LPRNet

روش	تعداد تصاویر	دقت (%)
LPRNet	۱۱۶۹۶	۹۵,۰
روش ما	۹۰۲	۹۲,۰

## ۵ جمع بندی

در این مقاله، یک الگوریتم جدید برای شناسایی شماره پلاک خودرو بر اساس شبکه‌های همگشتی عمیق و با رویکرد یادگیری با نظارت ضعیف پیشنهاد شد. یادگیری با نظارت ضعیف بدین معناست که برای هر تصویر پلاک در فاز آموزش تنها وجود کاراکترها مشخص می‌شود اما مکان آن‌ها مشخص نمی‌شود. برای هر کاراکتر، مشخصه‌ای که در تصاویر مثبت وجود دارد اما در تصاویر منفی وجود ندارد توسط شبکه همگشتی عمیق آموخته می‌شود. بنابراین، الگوریتم پیشنهادی قادر است برای تصاویر آزمون علاوه بر وجود هر کاراکتر، مختصات آنها را نیز مشخص نماید.

برای ارزیابی عملکرد الگوریتم پیشنهادی یک پایگاه داده شامل ۱۳۹۷ تصویر جمع‌آوری شد. نتایج ارزیابی 5-fold اثبات کرد که این الگوریتم می‌تواند به خوبی مشخصه‌های هر کاراکتر را بیاموزد به طوری که ۹۹,۱٪ از کاراکترها و ۹۵,۶٪ از پلاک‌ها را به درستی شناسایی نماید. موضوع پژوهش این مقاله تنها شناسایی شماره پلاک خودرو بوده است. در ادامه، آشکارسازی پلاک خودرو در تصویر با استفاده از رویکرد یادگیری با نظارت ضعیف مورد پژوهش قرار خواهد گرفت.

الگوریتم پیشنهادی در این مقاله برای کاراکترهایی که تعداد داده‌های آموزشی آنها کم باشد عملکرد قابل قبولی ندارد. به عبارت دیگر، نمی‌تواند ناحیه مربوط به کاراکتر مورد نظر را به درستی در تصاویر آموزشی مشخص نماید. به همین دلیل، یکی از کارهای آینده آن است که ناحیه برش خورده برای کاراکترهایی که فراوانی کمی دارند وارد شبکه شود تا تعیین محدوده این کاراکترها دچار خطا نگردد.

## مراجع

- [1] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.
- [2] R. O. Duda and P. E. Hart, "Use of the Hough transformation to detect lines and curves in pictures," *Commun. ACM*, vol. 15, pp. 11–15, 1972.

[31] S. Zherzdev and A. Gruzdev, "Lprnet: License plate recognition via deep neural networks," *arXiv preprint arXiv:10447*, 2018.



**محمدرضا محمدی** دوره‌های کارشناسی و کارشناسی ارشد خود را در دانشگاه صنعتی امیرکبیر گذرانده است و مدرک دکترای خود را از دانشگاه صنعتی شریف دریافت کرده است. وی در حال حاضر استادیار گروه هوش مصنوعی و رباتیکز دانشکده مهندسی کامپیوتر دانشگاه علم و صنعت ایران است. زمینه‌های پژوهشی و صنعتی مورد علاقه ایشان شامل پردازش تصویر، بینایی ماشین، هوش مصنوعی و یادگیری ماشین است.

- classifier ensemble," in *2015 Signal Processing and Intelligent Systems Conference (SPIS)*, 2015, pp. 48-52.
- [18] L. Zheng, X. He, B. Samali, and L. T. Yang, "An algorithm for accuracy enhancement of license plate recognition," *Journal of computer system sciences*, vol. 79, no. 2, pp. 245-255, 2013.
- [19] A. Abdussalam, S. Sun, M. Fu, H. Sun, and I. Khan, "License Plate Segmentation Method Using Deep Learning Techniques," in *International Conference On Signal And Information Processing, Networking And Computers*, 2018, pp. 58-65.
- [20] Y. Wang, X. Ban, J. Chen, B. Hu, and X. Yang, "License plate recognition based on SIFT feature," *Optik*, vol. 126, no. 21, pp. 2895-2901, 2015.
- [21] H. Chen, B. Hu, X. Yang, M. Yu, and J. Chen, "Chinese character recognition for LPR application," *Optik-International Journal for Light Electron Optics*, vol. 125, no. 18, pp. 5295-5302, 2014.
- [22] S. Ktata, T. Khadhraoui, F. Benzarti, and H. Amiri, "Tunisian License Plate Number Recognition," *Procedia Computer Science*, vol. 73, pp. 312-319, 2015.
- [23] V. H. Pham, P. Q. Dinh, and V. H. Nguyen, "CNN-Based Character Recognition for License Plate Recognition System," in *Asian Conference on Intelligent Information and Database Systems*, 2018, pp. 594-603.
- [24] P. Dhar, S. Guha, T. Biswas, and M. Z. Abedin, "A System Design for License Plate Recognition by Using Edge Detection and Convolution Neural Network," in *International Conference on Computer, Communication, Chemical, Material and Electronic Engineering 2018*, pp. 1-4.
- [25] M. Oquab, L. Bottou, I. Laptev, and J. Sivic, "Weakly supervised object recognition with convolutional neural networks," in *Proceedings of the Neural Information Processing Systems Conference*, 2014.
- [26] M. Oquab, L. Bottou, I. Laptev, and J. Sivic, "Is object localization for free?-weakly-supervised learning with convolutional neural networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 685-694.
- [27] Z. Jie, Y. Wei, X. Jin, J. Feng, and W. Liu, "Deep self-taught learning for weakly supervised object localization," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 1377-1385.
- [28] S. M. Pizer *et al.*, "Adaptive histogram equalization and its variations," *Computer vision, graphics, image processing*, vol. 39, no. 3, pp. 355-368, 1987.
- [29] G. Navarro, "A guided tour to approximate string matching," *ACM computing surveys*, vol. 33, no. 1, pp. 31-88, 2001.
- [30] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 248-255.