

بررسی کاهش خطای ناشی از شمارش تردد وسایل نقلیه در دوربین‌های پلاک خوان شهری توسط مدل‌های ریاضی (مطالعه موردی: شهر مشهد)

مقاله علمی - پژوهشی

محمدامین ابراهیم زاده*، گروه مهندسی عمران، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران

مهدیسه دباغی، گروه مهندسی عمران، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران

نوید نخعی، گروه مهندسی عمران، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران

*پست الکترونیکی نویسنده مسئول: m.ebrahimzadeh@iau.ir

دریافت: ۱۴۰۳/۰۹/۰۵ - پذیرش: ۱۴۰۴/۰۳/۰۱

صفحه ۲۱۳-۲۲۸

چکیده

یکی از مهم‌ترین ابزارهای سیستم‌های حمل و نقل هوشمند دوربین‌های کنترل سرعت می‌باشند و از متداولترین دوربین‌های موجود که در شبکه‌های شهری مورد استفاده قرار می‌گیرد، دوربین‌های پلاک خوان شهری است. بنابراین با توجه به اهمیت سنجش عملکرد دوربین‌های پلاک خوان شهری به منظور بررسی و شناخت صحیحی از وضعیت ترافیک شهرها، پژوهشی جهت شناسایی خطاهای موجود و سپس ارائه مدل ریاضی جهت کاهش خطای دوربین‌ها انجام شده است. در این پژوهش نمونه داده‌ها مربوط به اطلاعات ترافیکی ثبت شده توسط دو دوربین ترافیکی در شهر مشهد و در طول یک هفته اول خردادماه ۱۳۹۹ است. خطاهای شناسایی شده در این تحقیق شامل: خطای کل دوربین‌ها، خطای ثبت پلاک، خطای ثبت حرف پلاک، خطای ثبت کد شهر، خطای ثبت تخلف، زوج یا فرد بودن پلاک بودن می‌باشد. عوامل مختلفی بر بروز این خطاها تاثیر مستقیمی داشته است. از جمله موارد شماره دوربین، روز هفته، ساعت حرکت خودرو، سرعت حرکت خودرو، خط عبوری، تراکم یا حجم خودرو است. در این پژوهش در جهت شناسایی و کاهش خطاهای موجود، دو مدل کلاس‌بندی رگرسیون لجستیک و k نزدیک‌ترین همسایگی با توجه به پارامترهای موجود مورد بررسی و تحلیل قرار گرفته‌اند. در خطای کل موثرترین پارامتر سرعت بوده و مدل K نزدیک‌ترین همسایگی با دقت ۶۸ درصد بهترین کلاس‌بند برای خطای کل ثبت اطلاعات خودرو ارائه شده است.

واژه‌های کلیدی: دوربین ANPR، رگرسیون لجستیک، وضعیت ترافیک، مدل K

۱-مقدمه

خواهد بود. این افزایش جمعیت باعث افزایش تعداد خودروها به ۲,۵ میلیارد خواهد شد. امروزه در اغلب شهرهای پیشرفته دنیا، توسعه فناوری اطلاعات و ارتباطات را از طریق برنامه‌های جامع در قالب شهرهای هوشمند دنبال می‌نمایند که به کمک آن می‌توان به یکپارچگی اجزا شهر کمک کرده و به آن انسجام بخشید. یکی از کاربردهای فناوری در مدیریت شهری،

مدیریت هوشمند و نظارت بر ترافیک از الزامات اساسی توسعه شهرها محسوب می‌شود. گسترش الگوی شهرنشینی از سال ۱۹۵۰ به بعد و به دنبال آن افزایش شهرنشینی، وضعیت ترافیک در بسیاری از شهرها را چالش برانگیز شده است. براساس گزارش اخیر سازمان ملل متحد، رشد جمعیت و شهرنشینی تا سال ۲۰۵۰ با افزایش ۲,۹ میلیارد نفری روبه‌رو

پیشرفته، شماره پلاک خودروها را با دقت بالا ثبت می نماید. سپس توسط نرم افزارهای مربوطه، پردازش داده می شود [Albayrak, 2015]. تحقیقی سیستم های ANPR اراده شده است. در این پژوهش به بررسی ویژگی های فیزیکی پلاک ها و نیز ویژگی های دوربین ها و تاثیر آنها بر میزان خطای سیستم های پرداخته شده است. در این پژوهش میزان نور ثابت، ارتفاع دوربین، و زاویه دید مناسب به عنوان پارامتر های بسیار موثر در میزان خطا در دوربین ها در نظر گرفته شده است. همچنین این پژوهش نشان می دهد میزان خطا در شب بسیار بیشتر از میزان خطا در روز و روشنایی است [Robinson et al, 2019]. آمار تردد وسایل نقلیه از اساسی ترین و مهم ترین آمارهای مورد نیاز در مدیریت حمل و نقل است. امروزه تردد شمارها به عنوان یکی از روش های نوین اخذ اطلاعات برخط لحظه های در مدیریت ترافیک مطرح است. با وجود اهمیت نصب و استقرار چنین تجهیزاتی، مهمترین مسئله و دغدغه، تعیین مکان بهینه نصب آنها است تا بتوان به آمار قابل استنادی به منظور مدیریت راهها دست یافت؛ بنابراین هدف اصلی این پژوهش، ارائه مدلی برای مکانیابی تردد شمار در راه های برون شهری به منظور استخراج مقدار بهینه AADT خواهد بود [فائزی و همکاران، ۱۳۹۷]. سرعت وسایل نقلیه یکی از عوامل مهم تعیین کننده شدت و میزان جراحت و تصادفات منجر به فوت سرنشینان وسایل نقلیه است. به منظور ارتقای ایمنی در جاده ها و راه های برون شهری، مسئولان حوزه حمل و نقل، تصمیم بر ایجاد محدودیت های سرعت گرفته اند. در این مطالعه با استفاده از دو روش از مطالعات قبل و بعد، روش تجربی بیز و گروه مقایسه، داده های جمع آوری شده از تردد شمارهای آزادراه قزوین-رشت و بزرگراه بم-کرمان مورد ارزیابی قرار می گیرد تا میزان تاثیر این دوربین ها در کاهش تعداد تخلفات سرعت سنجیده شود. نتایج روش بیز برای دوربین کنترل سرعت در محور رودبار-رستم آباد نشان دهنده کاهش ۸۹/۵۱ و نتایج گروه مقایسه آن نشان دهنده کاهش ۸۱ تا ۹۵ درصدی تخلفات سرعت و نتایج روش بیز برای محور جیرفت-ماهان نشان دهنده کاهش ۹۴ درصدی تخلفات سرعت و نتایج روش گروه مقایسه نشان دهنده کاهش تقریبی ۸۰ تا ۸۸ درصدی تخلفات سرعت در محدوده دوربین کنترل سرعت است. پس در مجموع می توان گفت که دوربین های کنترل سرعت در این محورها دارای اثرات مثبتی

سیستم های هوشمند حمل و نقلی است. سیستم های حمل و نقل هوشمند با بکارگیری تکنولوژی های نوین از قبیل الکترونیک، ارتباطات و سیستم های کنترل رایانه ای و تلفیق آنها با علوم مهندسی ترافیک و برنامه ریزی، ابزارهای لازم برای کنترل و مدیریت را فراهم ساخته اند. به طور کلی تعریفی از حمل و نقل هوشمند در سال ۱۹۹۸ توسط انجمن حمل و نقل هوشمند آمریکا ارائه شد که به صورت عام مورد قبول قرار گرفت که به بدین شرح است [Schreieck, 2017]. مردم از تکنولوژی در حمل و نقل برای صرفه جویی در پول و وقت در زندگی روزمره استفاده می کنند. تعریف رسمی تری نیز توسط حمل و نقل آمریکا در آوریل سال ۱۹۹۹ نیز منتشر شد که بدین شرح است: سیستم حمل و نقل هوشمند اطلاعات مربوط به جا به جایی مسافر و کالا را جمع آوری، نگهداری، پردازش و توزیع می نماید یکی از مهم ترین ابزارهای سیستم های حمل و نقل هوشمند، دوربین های کنترل ترافیک می باشند. این دوربین ها به عنوان یک سیستم نظارتی، کاربردهای متعددی در حمل و نقل دارند که شامل: کنترل ترافیک، ثبت تخلفات، نظارت بر رعایت قوانین راهنمایی و رانندگی از جمله کنترل سرعت، اندازه گیری زمان سفر می شود. استفاده از دوربین ها سبب افزایش ایمنی در سیستم حمل و نقل درون شهری می شوند. همچنین به کمک این دوربین ها می توان موضوعاتی از جمله شرایط جوی و گزارش تصادفات را نیز به ثبت رساند. استفاده از پردازش تصویر برای نظارت ترافیک در اواسط سال ۱۹۷۰ میلادی آغاز شد. سخت افزار و الگوریتم های مورد استفاده برای پارامترهای ترافیک باعث بهبود بزرگی در طول سال های گذشته شده است. در مدل سازی شبکه های ترافیک، پارامترهایی نظیر تعداد خودرو، سرعت، مسیرهای خط سیر و طبقه بندی و همچنین فاصله متوالی بین وسایل نقلیه و پارامترهای دیگر که جزء عوامل مهم ترافیک می باشند و نقش کلیدی دارند با این سخت افزار امکان پذیر می باشند. تشخیص خودرو برای برنامه های کاربردی نظارت در بزرگراه ها و برنامه ریزی ترافیک شهری بسیار مهم است [Gaber et al, 2018]. یکی از متداول ترین دوربین های موجود که در شبکه های شهری مورد استفاده قرار می گیرد، دوربین های پلاک خوان شهری است. این سامانه کاربردهای بسیاری در سامانه های نظارت بر ترافیک روزانه و سامانه های کنترل عوارض جاده ای دارند. این نوع سیستم ها تصاویر خودروها را ضبط کرده و با استفاده از الگوریتم های

دسته‌ای می‌باشد. در رگرسیون لجستیک بجای پیش‌بینی خود متغیر پاسخ، احتمال هر یک از برآمدهای ممکن متغیر پاسخ پیش‌بینی می‌شود [Chitarron.2002].

۲-۱- روش K نزدیک‌ترین همسایگی

الگوریتم K نزدیک‌ترین همسایگی یک روش ناپارامتریک و یکی از ساده‌ترین روش‌های دسته‌بندی در یادگیری ماشین است. این روش بر این پیش‌فرض بنا شده است که داده‌های مشابه متعلق به کلاس یکسانی هستند. لذا دسته‌بندی یک مورد بر مبنای رای اکثریت نقاط همسایگی انجام می‌شود. به‌طور کلی، این روش برای طبقه‌بندی یک عنصر در یک مجموعه است که این کار را براساس نزدیکترین خصوصیات سایر اعضای موجود در مجموعه (نمونه‌های آموزشی) انجام می‌دهد. این روش دارای یکی از ساده‌ترین الگوریتم‌های آموزش برای پیش‌بینی داده‌ها است و بر خلاف روش‌های دیگر از هیچ تابع ریاضیاتی از پیش مشخص شده‌ای، برای تخمین عناصر هدف استفاده نمی‌کند. الگوریتم K نزدیک‌ترین همسایگی به گروه موسوم به الگوریتم‌های یادگیرنده تنبل تعلق دارد و داده‌ها را به‌گونه‌ای ذخیره می‌نماید که تا زمانی که دستوری برای برآورد جدید نباشد، فرآیند آموزش روی این داده‌ها انجام نخواهد شد. این الگوریتم یک متد ناپارامتری است که برای طبقه‌بندی آماری و رگرسیون استفاده می‌شود. در هر دو حالت کی شامل نزدیک‌ترین مثال آموزشی در فضای داده ای می‌باشد و خروجی آن بسته به نوع مورد استفاده در طبقه بندی و رگرسیون متغیر است. [حاجلو و همکاران، ۱۳۹۴].

۳- مطالعه موردی

مشهد کلان‌شهری در شمال شرقی ایران و مرکز استان خراسان رضوی است. این شهر با ۳۵۱ کیلومتر مربع مساحت، دومین شهر پهناور ایران پس از تهران است. براساس سرشماری عمومی نفوس و مسکن سال ۱۳۹۵ مشهد با ۳,۰۰۱,۱۸۴ تن جمعیت، دومین شهر پرجمعیت ایران پس از تهران و نود و پنجمین شهر پر جمعیت دنیا به‌شمار می‌رود. این شهر به واسطه وجود حرم علی بن موسی الرضا، هشتمین امام مذهب شیعه، مهم‌ترین قطب زیارتی و کانون عمده گردشگری مذهبی و زیارت در ایران شناخته شده و سالانه پذیرای بیش از ۲۷ میلیون زائر از داخل و دو میلیون زائر از خارج از کشور است.

در کاهش چشم‌گیر تخلیفات سرعت در محدوده نصب دوربین با استفاده از روش‌های مذکور است [امیری و همکاران، ۱۳۹۹]. اعمال روش‌های گوناگون مدیریت سرعت در راه‌ها و کنترل کردن سرعت حرکت وسایل نقلیه در محدوده سرعت ایمن، یکی از روش‌های مطرح در پیشگیری از وقوع بسیاری از تصادفات رانندگی در جاده‌ها است. یکی از مهم‌ترین و کارآمدترین روش‌های جلوگیری از افزایش سرعت غیر مجاز در جاده‌ها و در پی آن پایین آمدن میزان تصادفات، بکارگیری ابزارهای کنترل سرعت در جاده‌ها می‌باشد. با توجه به هزینه بالای بکارگیری سیستم‌های هوشمند کنترل سرعت و محدودیت در بودجه مالی پروژه‌ها و مجموعه محدودیت‌های دیگر که براساس شرایط هر پروژه در نظر گرفته می‌شود. در این مقاله سعی بر آن بوده تا با توجه به محدودیت‌های نصب و بکارگیری دوربین‌های کنترل سرعت نسبت به ارائه یک مدل بهینه‌سازی با هدف بهینه‌سازی پوشش آنها در یک مسیر اقدام شود. محدودیت‌های مدل بهینه‌سازی ارائه شده بنا بر مشخصات فیزیکی مسیر، محدودیت‌های دوربین‌ها، سوابق تصادفات در مقطع و سایر موارد است. اعتبار مدل ارائه شده بر روی مقطعی از جاده کرج - چالوس حد فاصل میدان امیرکبیر تا شهرستان گچسار به طول ۶۰ کیلومتر انجام شد و نتایج حاصل از مکانیابی در حالات مختلف به ازای میزان بودجه مالی در دسترس مورد تحلیل قرار گرفت [امیربها و همکاران، ۱۳۹۶].

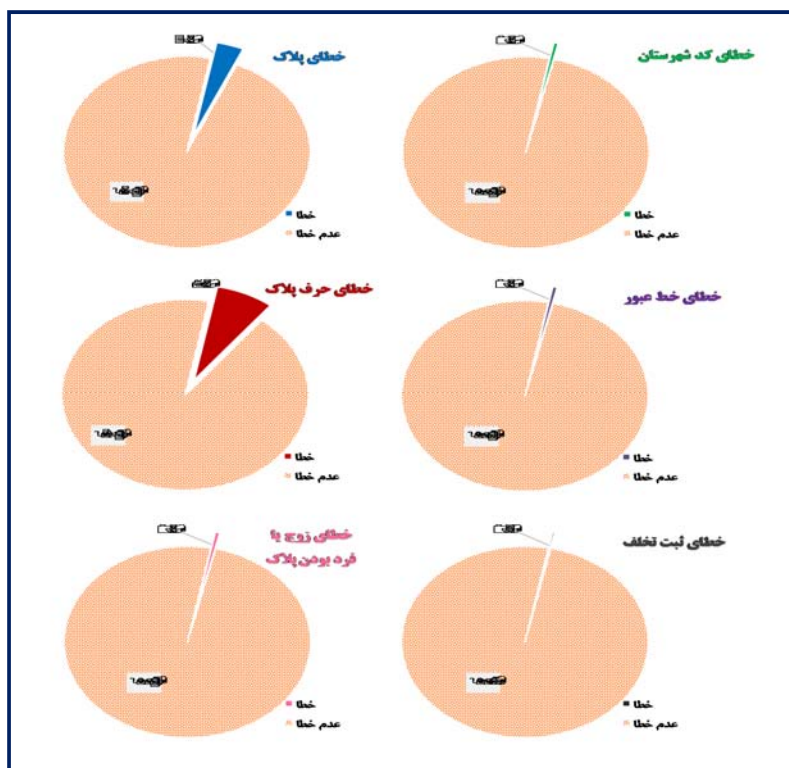
۲- روش شناسی

این تحقیق به طراحی مدل ریاضی کاهش خطای ناشی از شمارش تردد وسایل نقلیه در دوربین‌های پلاک‌خوان شهری می‌پردازد و از طریق یافته‌های آن سعی خواهد شد راه‌کارهایی در جهت اصلاح و رفع خطاهای شناسایی شده ارائه شود، لذا نوع این تحقیق از حیث هدف، کاربردی قلمداد می‌شود. لازم به ذکر می‌باشد که یافته‌های تحقیق کاربردی به میزان بسیار زیادی قائم به زمان و مکان هستند. رگرسیون لجستیک یکی از مهم‌ترین مدل‌های دسته‌بندی می‌باشد و تنها تفاوت آن با رگرسیون خطی در مقیاس اندازه‌گیری متغیر هدف یا متغیر پاسخ است. به‌طوری که در رگرسیون خطی متغیر هدف عددی یا پیوسته (صفت کمی با مقیاس اندازه‌گیری فاصله‌ای یا نسبی) در نظر گرفته می‌شود و سعی می‌کند رابطه بین متغیرهای پیش‌بینی‌کننده و متغیر هدف را در قالب یک تابع برآورد کند. در حالی که در یک مسئله دسته‌بندی، متغیر هدف به‌صورت

شناخت خطاهای موجود در عملکرد دوربین‌های ANPR و تفکیک آن‌ها

بررسی وضعیت خطاها در ثبت پلاک و خط عبوری خودروها (عدم وضوح و شفافیت تصاویر دوربین‌ها برای خوانش پلاک خودرو) هدف اصلی تحقیق حاضر است. لذا در این بخش توصیف کلی از انواع خطای موجود در اطلاعات ترافیکی دوربین‌ها ارائه می‌گردد.

باتوجه به جمعیت فعلی و جمعیت زائر و مهاجر موجود در این شهر و به منظور مدیریت بهتر و کارآمدتر شهری به بررسی میزان خطا در برداشت و خوانش دوربین‌های پلاک خوان شهری، در این کلان شهر پرداخته‌ایم.



شکل شماره ۱. نمودار دایره‌ای انواع خطای اطلاعات ترافیکی ثبت شده توسط دوربین‌ها

۴-مدلسازی و تحلیل خروجی‌ها

مذکور از طریق انجام فرآیند داده‌کاوی نقشه‌ی اساسی در مدیریت این دوربین‌ها و افزایش کارایی آنها خواهد داشت.

۴-۱- ورود اطلاعات و داده‌ها

با توجه به هدف تحقیق حاضر، داده‌های مورد استفاده در این تحقیق مربوط به اطلاعات ترافیکی ثبت شده توسط دو دوربین ترافیکی به شماره‌های ۱۳، به آدرس مشهد - بزرگراه شهید کلاتری، درب شرقی دانشگاه فردوسی، جنوب - شرقی و دوربین شماره ۱۴، به آدرس مشهد، بزرگراه شهید کلاتری، بلوار صدا و سیما، شمال - غربی در طول یک هفته اول خردادماه ۱۳۹۹ می‌باشد.

پژوهش حاضر که با استفاده از متدولوژی CRISP-DM انجام گرفته است شامل مراحل زیر است. با توجه به اینکه هدف اصلی تحقیق حاضر، شناخت خطاهای موجود در عملکرد دوربین‌های ANPR و بررسی تأثیر متغیرهای مستقل نظیر: اطلاعات خودرو، زمان، سرعت خودرو و خط عبوری بر میزان خطای دوربین‌های ANPR می‌باشد، هر موضوعی مرتبط با دوربین‌های ترافیکی ANPR بهره‌مند از نتایج این تحقیق بوده و در دامنه اهداف تجاری-عملیاتی این طرح داده‌کاوی قرار می‌گیرد. شناسایی خطاهای موجود، عوامل موثر بر این خطاها، کشف ارتباط بین متغیرهای ممکن و بررسی الگوی رفتاری داده‌های ثبتی دوربین‌های

اولیه از تجمیع و چسباندن ۳۷۰۲ فایل اکسل شامل صد رکورد از اطلاعات ثبتی دوربین شماره ۱۳ و ۴۲۵۷ فایل اکسل شامل صد رکورد از اطلاعات ثبتی دوربین شماره ۱۴ با استفاده از کدنویسی در ویژوال بیسیک به دست آمده است.

که شامل اطلاعات پلاک خودرو، تخلف ترافیکی، زمان دقیق، سرعت و تعداد خودرو است. برای انجام این تحقیق، ابتدا با مطالعه و بررسی داده های موجود، شاخص های مؤثر بر خطای دوربین های ترافیکی شناسایی و بانک اطلاعاتی شامل ۱۲ ستون (ویژگی) و ۷۸۸۲۰۲ سطر (رکورد) تشکیل شد. بانک اطلاعاتی

جدول ۱. توصیف متغیرهای تحقیق

متغیرها	نوع متغیر	نوع/مقیاس اندازه گیری	نوع داده
شماره دوربین	کیفی	اسمی	گسسته
آدرس دوربین	کیفی	اسمی	گسسته
تاریخ ثبت تصویر	کمی	فاصله ای	گسسته
روز هفته	کیفی	ترتیبی	گسسته
زمان دقیق ثبت تصویر	کمی	فاصله ای	پیوسته
سرعت خودرو	کمی	نسبی	پیوسته
پلاک خودرو	کیفی	اسمی	گسسته
حرف پلاک	کیفی	اسمی	گسسته
شهر پلاک	کیفی	اسمی	گسسته
خط عبوری	کیفی	اسمی	گسسته
تخلف	کیفی	اسمی	گسسته
زوج یا فرد بودن پلاک	کیفی	اسمی	گسسته

۴-۲- تفکیک داده ها بر اساس نمونه متغیرها

دوربین های ترافیکی: در پژوهش حاضر، داده های دو دوربین ترافیکی به آدرس های «بزرگراه کلاتری - درب شرقی دانشگاه، جنوب-شرقی» و «بزرگراه کلاتری - صدا و سیما، شمال-غربی» در طول یک هفته اول خرداد مورد بررسی قرار گرفته است.

داده های تحقیق در قالب متغیرهای دوربین ترافیکی، روزهای هفته، زمان تصویر برداری، سرعت حرکت خودرو مورد تفکیک و بررسی قرار گرفته، که در ذیل به شرح هر یک پرداخته شده است. همچنین دو پارامتر خط عبوری و حجم ترافیک نیز بر خطای دوربین ها تاثیر به سزایی دارند.

جدول ۲. توزیع فراوانی شماره / آدرس دوربین های ترافیکی

شماره دوربین	آدرس دوربین	تعداد	درصد
۱۳	بزرگراه کلاتری، درب شرقی دانشگاه	۳۶۲۶۰۰	۴۶
۱۴	بزرگراه کلاتری، صدا و سیما	۴۲۵۶۰۲	۵۴
جمع		۷۸۸۲۰۲	۱۰۰

روزهای هفته: برای بررسی اثر روزهای هفته بر عملکرد دوربین‌های ترافیکی، داده‌های ترافیکی مرتبط با تمامی روزهای هفته اول خردادماه ۱۳۹۹ مورد بررسی قرار گرفت. در ادامه نحوه توزیع فراوانی داده‌های ترافیکی در روزهای هفته گزارش می‌گردد.

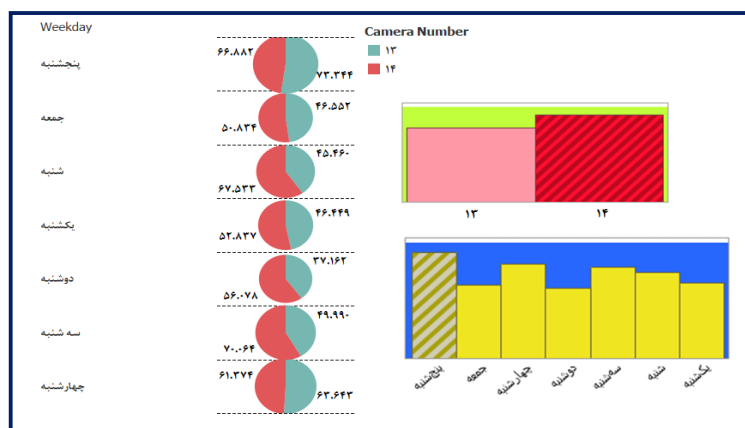
نتایج جدول فوق نشان می‌دهد، ۵۴ درصد داده‌های ترافیکی (ثبت لحظه‌ای پلاک‌های خودروهای ترددی از بزرگراه کلاتری) مربوط به دوربین شماره ۱۴ (به آدرس بزرگراه کلاتری صدا و سیما) و ۴۶ درصد داده‌های ترافیکی مربوط به دوربین شماره ۱۳ (به آدرس بزرگراه کلاتری درب شرقی دانشگاه) می‌باشد.

جدول ۳. توزیع فراوانی روزهای هفته

تاریخ	روزهای هفته	تعداد	درصد
۱۳۹۹/۰۳/۰۱	پنجشنبه	۱۴۰۲۲۶	۱۸
۱۳۹۹/۰۳/۰۲	جمعه	۹۷۳۸۶	۱۲
۱۳۹۹/۰۳/۰۳	شنبه	۱۱۲۹۹۳	۱۴
۱۳۹۹/۰۳/۰۴	یکشنبه	۹۹۲۸۶	۱۳
۱۳۹۹/۰۳/۰۵	دوشنبه	۹۳۲۴۰	۱۲
۱۳۹۹/۰۳/۰۶	سه‌شنبه	۱۲۰۰۵۴	۱۵
۱۳۹۹/۰۳/۰۷	چهارشنبه	۱۲۵۰۱۷	۱۶
	جمع	۷۸۸۲۰۲	۱۰۰

جداول شماره ۲ و ۳ نمودارهای ستونی و دایره‌ای تهیه شده که در شکل شماره ۲ آمده است. این شکل نتایج بدست آمده در جداول فوق را به‌خوبی نمایان می‌سازد.

همان‌طور که نتایج مندرج در جدول فوق نشان می‌دهد، بیشترین فراوانی با ۱۸ درصد مربوط به داده‌های ترافیکی روز پنجشنبه می‌باشد. جهت برازش یک مدل گرافیکی به نتایج



شکل ۲. توزیع فراوانی شماره دوربین ترافیکی در کل و به تفکیک روزهای هفته

قالب متغیری گسسته برحسب ساعت (ساعت زمان ثبت شده) در کل و به تفکیک روزهای هفته تحلیل می‌شود.

زمان تصویربرداری: یکی دیگر از متغیرهای مورد بررسی زمان ثبت تصاویر (ساعت، دقیقه، ثانیه) می‌باشد که در ادامه توزیع فراوانی متغیر زمان در قالب متغیری پیوسته برحسب ثانیه و در

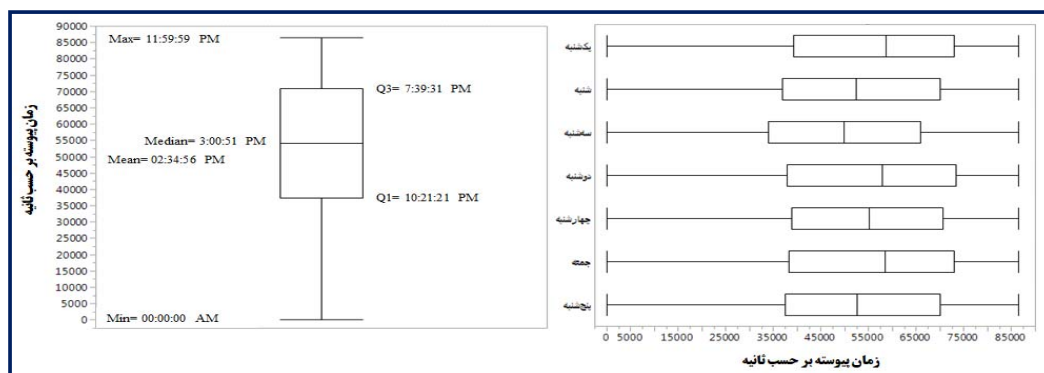
$$\text{Continuous Time Variable} = \text{Hour} \times 3600 + \text{Minute} \times 60 + \text{Second}$$

جدول ۴: شاخص‌های توصیفی متغیر زمان تصویربرداری

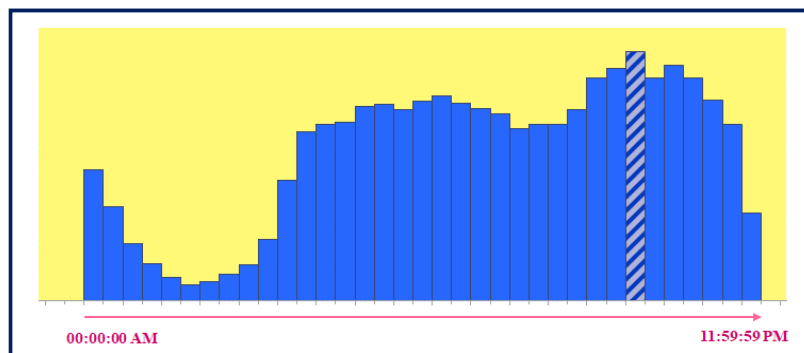
متغیر	تعداد	میانگین	میانه	کمترین امتیاز	بیشترین امتیاز	انحراف معیار
زمان تصویربرداری (تعداد ثانیه)	۷۸۸۲۰۲	۵۲۴۹۶	۵۴۰۵۱	۰	۸۶۳۹۹	۲۱۶۷۹
ساعت تصویربرداری	۷۸۸۲۰۲	۱۴/۱	۱۵	۰	۲۳	۶

با توجه به شکل شماره ۱ بیشترین فراوانی تصاویر (داده‌ها) مربوط به ساعات ۱۸ الی ۲۲ می‌باشد. با توجه به نمودار سمت راست شکل فوق می‌توان مشاهده نمود که توزیع فراوانی زمان تصاویر (تصویربرداری‌ها) در روزهای مختلف تقریباً یکسان است و اختلاف اندک و ناچیزی دارند.

نتایج این جدول حاکی از آن است که میانگین زمان تصویربرداری ساعت ۲ و ۳۴ دقیقه و ۵۶ ثانیه می‌باشد. ۵۰ درصد تصاویر مربوط به قبل از ساعت ۱۵ و ۵۱ ثانیه و ۵۰ درصد دیگر به بعد از این ساعت تعلق دارد. در نمودارهای ادامه این نتایج به‌خوبی نمایش داده شده‌اند. به‌طوری‌که،

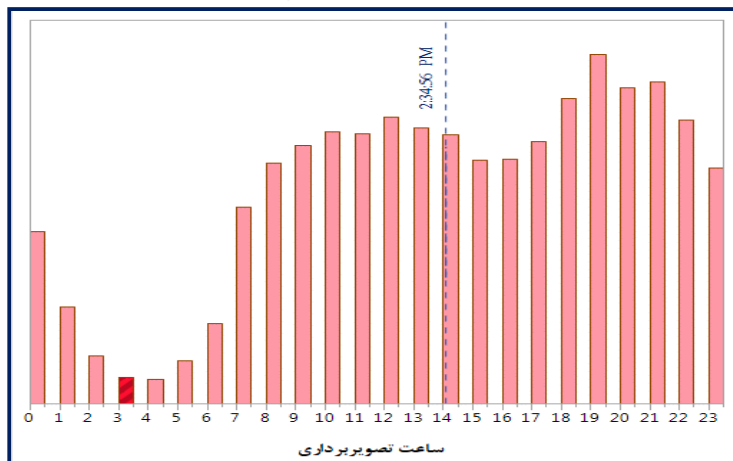


شکل ۳. نمودار جعبه‌ای توزیع فراوانی زمان ثبت تصاویر ترافیکی در کل و به تفکیک روزهای هفته

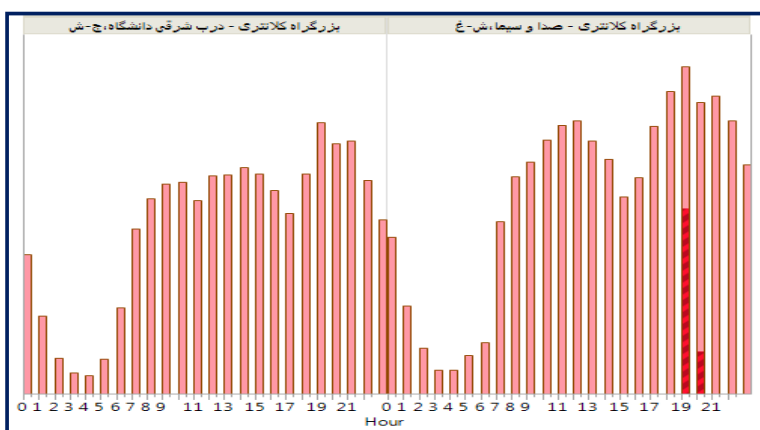


شکل ۴. نمودار هیستوگرام زمان ثبت تصاویر ترافیکی (تصویربرداری)

در ادامه نمودارهای مربوط به ساعت تصاویر یا داده‌های ثبتی را ارائه و توضیح می‌دهیم.



شکل ۵. نمودار ستونی ساعت ثبت تصاویر ترافیکی (ساعت تصویربرداری)



شکل ۶. نمودار ستونی مقایسه‌ای ساعت تصویربرداری دو دوربین ترافیکی

با توجه به نمودار شکل شماره ۵ می‌توان مشاهده نمود که ساعت تصاویر ثبت شده‌ی دو دوربین تقریباً هم توزیع هستند و اختلاف اندکی بین دوربین‌ها مشاهده می‌شود.

سرعت خودرو: سرعت خودرو یکی دیگر از متغیرهای مورد بررسی می‌باشد، که همزمان با تصویربرداری پلاک خودروهای عبوری از بلوار کلاتری توسط دوربین‌های شماره ۱۳ و ۱۴ ثبت می‌گردد.

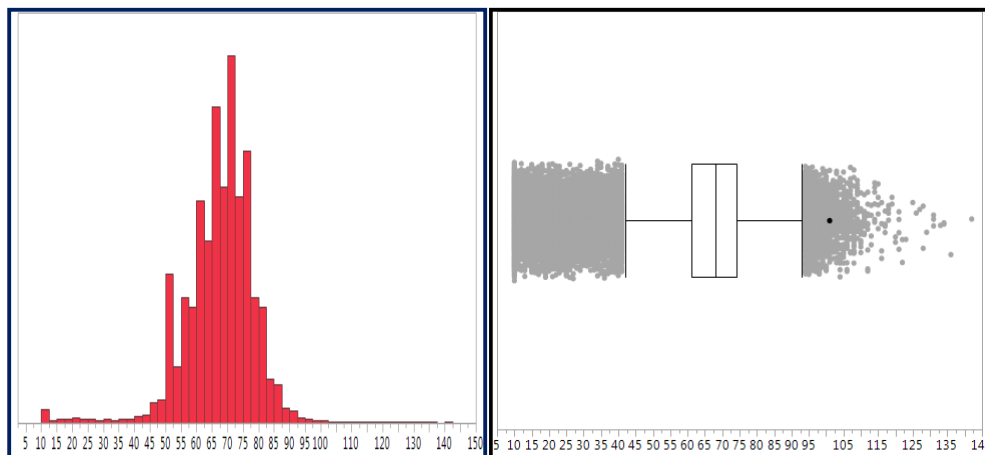
جدول ۵. شاخص‌های توصیفی متغیر سرعت خودرو

متغیر	تعداد	میانگین	میانه	کمترین امتیاز	بیشترین امتیاز	انحراف معیار
سرعت خودرو (کیلومتر بر ساعت)	۷۸۸۲۰۱	۶۷/۱	۶۸	۱۰	۱۴۲	۱۰/۹

طبقه‌بندی سرعت	کمتر از ۲۰	بین ۲۰ تا ۴۰	بین ۴۰ تا ۶۰	بین ۶۰ تا ۸۰	بین ۸۰ تا ۱۰۰	بیشتر از ۱۰۰
تعداد	۶۰۵۵	۵۲۴۲	۱۴۴۳۸۹	۵۴۵۱۳۶	۵۲۷۷۷	۳۴۶۰۲
درصد فراوانی	٪۱	٪۱	٪۱۸	٪۶۹	٪۷	٪۴

بر ساعت داشته‌اند. سرعت ۴ درصد خودروها بیشتر از ۱۰۰ کیلومتر بر ساعت می‌باشد. سرعت ۲ درصد خودروها کمتر از ۴۰ کیلومتر بر ساعت است. نتایج مذکور در نمودارهای هیستوگرام و جعبه‌ای رسم شده در اشکال ۷ و ۸ به خوبی به تصویر کشیده شده‌اند.

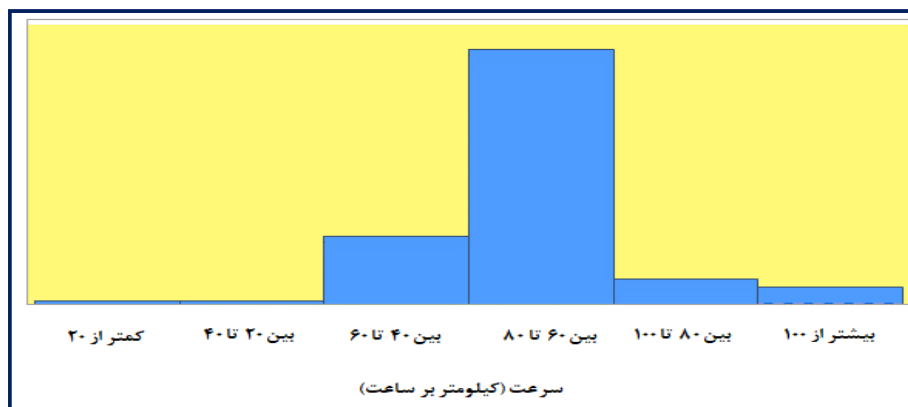
همان‌طور که جدول فوق نشان می‌دهد، میانگین و انحراف‌معیار به دست آمده برای سرعت حرکت خودروهای تصویربرداری شده به ترتیب ۶۷/۱ و ۱۰/۹ کیلومتر بر ساعت است. کمترین سرعت ثبت شده ۱۰ کیلومتر بر ساعت و بیشترین سرعت ۱۴۲ کیلومتر بر ساعت می‌باشد. سرعت ۵۰ درصد خودروها کمتر یا مساوی ۶۸ کیلومتر بر ساعت و سرعت ۵۰ درصد دیگر حداقل ۶۸ کیلومتر بر ساعت بوده است. ۶۹ درصد خودروها سرعتی بین ۶۰ تا ۸۰ کیلومتر



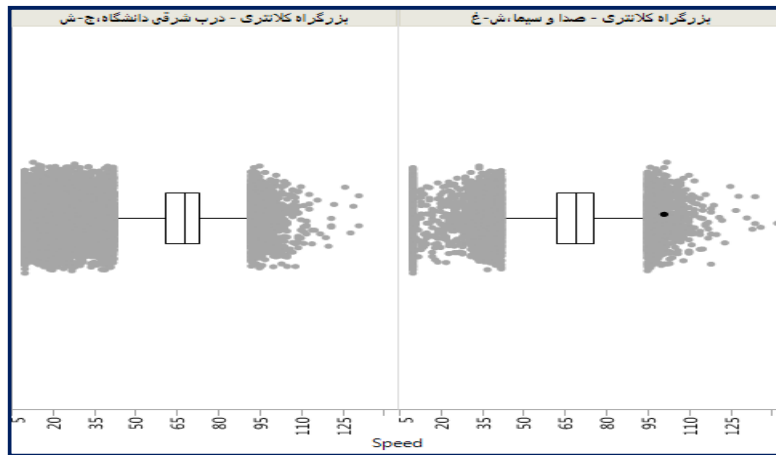
شکل ۷. نمودار هیستوگرام (سمت راست) و نمودار جعبه‌ای (سمت چپ) سرعت خودروهای تصویربرداری شده

بین داده‌های ثبتی دو دوربین ترافیکی مورد بررسی، مشاهده می‌شود که توزیع سرعت حرکت خودروها در دو طرف بلوار کلانتری تقریباً یکسان است.

در ادامه نمودار هیستوگرام سرعت خودرو در قالب متغیری ترتیبی (رتبه‌ای) ارائه می‌گردد. در هفته اول خرداد و در بلوار کلانتری ۸۷ درصد ماشین‌ها سرعتی بین ۴۰ تا ۸۰ کیلومتر بر ساعت داشته‌اند. در ادامه با مقایسه متغیر سرعت

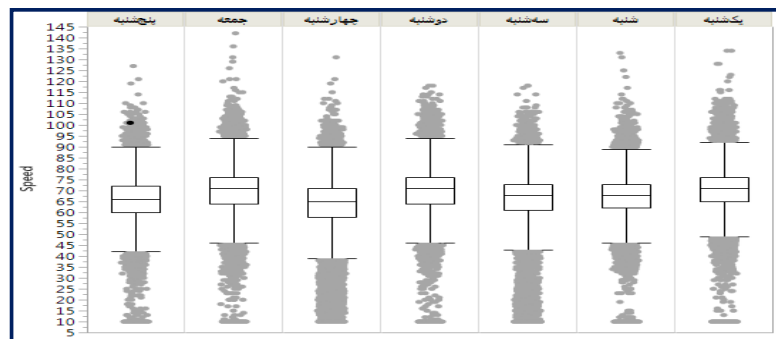


شکل ۸. نمودار هیستوگرام سرعت خودروهای تصویربرداری شده در قالب متغیری با مقیاس رتبه‌ای



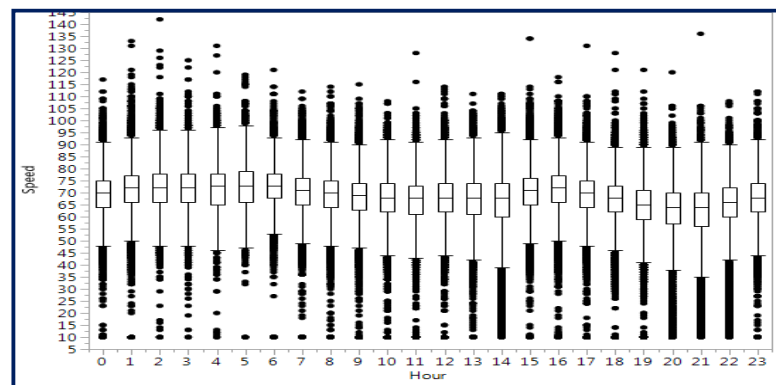
شکل ۹. نمودار جعبه‌ای سرعت خودروهای تصویربرداری شده به تفکیک دو دوربین ترافیکی ۱۳ و ۱۴

همان‌طور که شکل شماره ۱۰ نشان می‌دهد، توزیع فراوانی سرعت در روزهای مختلف هفته تقریباً یکسان است.



شکل ۱۰. نمودار جعبه‌ای سرعت خودروهای تصویربرداری شده به تفکیک روزهای هفته

رفتار خودروها از نظر سرعت در ساعات مختلف روز نیز تقریباً یکسان است.



شکل ۱۱. نمودار جعبه‌ای سرعت خودروهای تصویربرداری شده به تفکیک ساعات مختلف روز

بررسی میزان تأثیر زمان و سایر متغیرهای ترافیکی بر خطای دوربین‌ها

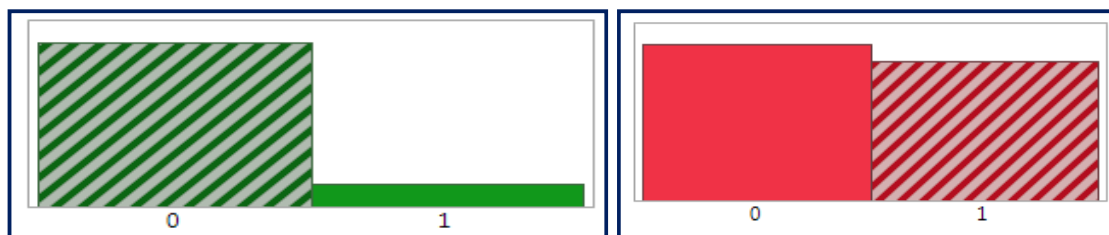
مرحله اول) متوازن کردن داده: متغیرهای زمان شامل روزهای هفته و ساعات مختلف شبانه‌روز است. در این بخش تأثیر متغیرهای زمانی و سایر متغیرها شامل سرعت، حجم خودرو در ساعت، شماره دوربین، خطوط عبوری خودرو بر خطای

داده‌های متوازنی که بر مبنای متغیر خطای کل و با استفاده از الگوریتم SMOTE ساخته و تولید گردید، شامل ۷۹۴۳۴۱ رکورد می‌باشد.

دوربین‌ها بررسی می‌شود. برای این منظور ابتدا داده‌های غیر متوازن بر اساس خطای کل در ثبت اطلاعات (مشاهده هریک از انواع خطاهای ثبت پلاک، ثبت حرف پلاک، ثبت کد شهر، ثبت تخلف، ثبت زوج یا فرد بودن پلاک) و سپس به تفکیک هر یک از خطاها با استفاده از الگوریتم SMOTE متوازن می‌گردد.

جدول ۶. توزیع فراوانی خطای کل ثبت اطلاعات خودرو توسط دوربین‌های ترافیکی در داده‌های متوازن

خطا	کد	تعداد	درصد
ثبت کامل اطلاعات خودرو	عدم خطا=۰	۶۹۴۷۵۰	۸۸
	خطا=۱	۹۳۴۵۲	۱۲
جمع			
۷۸۸۲۰۲			
ثبت کامل اطلاعات خودرو	عدم خطا=۰	۴۲۰۵۳۳	۵۳
	خطا=۱	۳۷۳۸۰۸	۴۷
جمع			
۷۹۴۳۴۱			



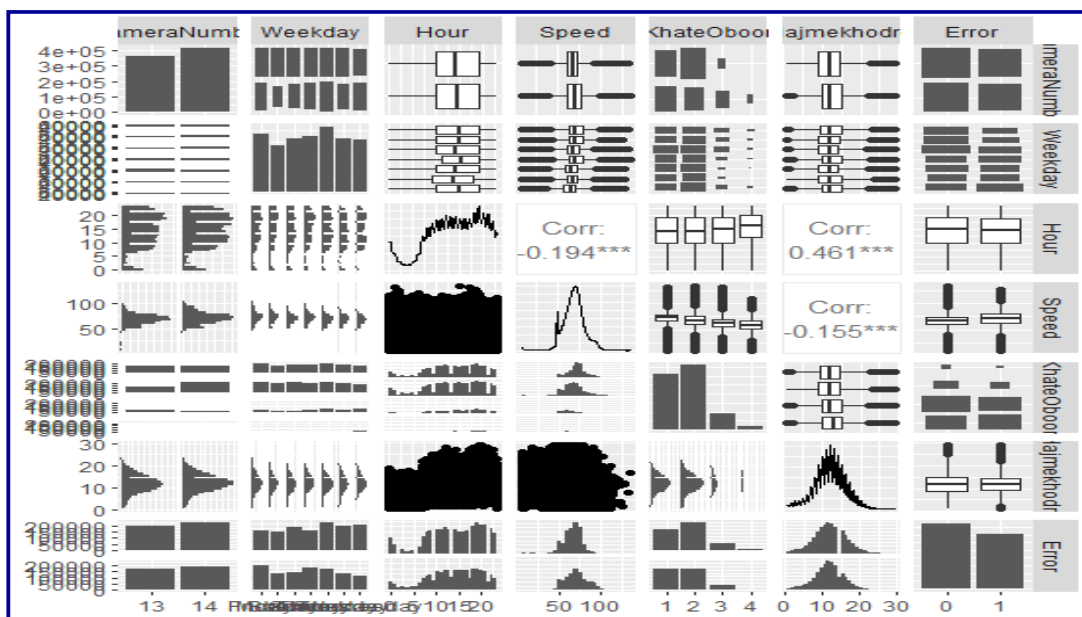
شکل ۱۲. نمودار ستونی توزیع فراوانی خطای کل ثبت اطلاعات خودرو توسط دوربین‌های ترافیکی در داده‌های متوازن (سمت راست) و غیر متوازن (سمت چپ)

نتایج مدل‌ها

مرحله سوم) ورود اطلاعات به مدل‌ها و تحلیل نتایج: نتایج چهار الگوریتم کلاس‌بندی (رگرسیون لجستیک، k نزدیک‌ترین همسایگی) مورد بررسی قرار گرفته است.

الف) روش K نزدیک‌ترین همسایگی: برای بررسی تأثیرگذاری متغیرهای مستقل فوق‌الذکر بر وقوع خطای کل (هرگونه خطا در ثبت اطلاعات خودرو) ابتدا از روش k نزدیک‌ترین همسایگی استفاده گردید.

ابتدا در ماتریس نمودار پراکنش که در شکل شماره ۱۳ ذیل ارائه گردیده است، رابطه بین هر دو متغیر مذکور مورد مطالعه قرار می‌گیرد. با توجه به اینکه در ۱۶۱۴۱ رکورد از داده‌های متوازن خطای ثبت خط عبور اتفاق افتاده و دقیقاً در تمامی این موارد یکی دیگر از سایر انواع خطا نیز به وقوع پیوسته است، لذا برای اجتناب از ایجاد اربیبی در نتایج مدل‌های برازش شده روی داده‌ها این تعداد رکورد از بانک داده حذف می‌گردد.



شکل ۱۳. ماتریس نمودار پراکنش متغیرهای حاضر در مدل‌های کلاس‌بندی مورد استفاده برای متغیر خطا در ثبت اطلاعات خودرو

جدول ۷. ماتریس انطباقی داده‌های آموزشی روش k نزدیک‌ترین همسایگی

واقعی		ماتریس انطباقی داده‌های آزمایشی	
خطا	عدم خطا	خطا	پیش‌بینی شده
۱۹۶۹۸	۵۲۱۰۸	خطا	پیش‌بینی شده
۱۰۶۲۷۹	۵۵۳۷۵	عدم خطا	

Accuracy = 0.678

دقت مدل رگرسیون لجستیک برآزش یافته روی داده‌ها نسبتاً پایین است. تمام متغیرهای مستقل درون مدل شامل روز هفته، خط عبور، حجم تردد، ساعت حرکت، سرعت حرکت و شماره دوربین بر احتمال وقوع خطای کل در ثبت اطلاعات خودرو توسط دوربین‌های ترافیکی تأثیرگذارند. نمودارهای پراکنش و نمودارهای موزائیکی رسم شده نحوه تأثیرگذاری متغیرهای مستقل را نشان می‌دهد. همچنین ضریب تشخیص هر یک از متغیرهای مستقل در جدول شماره ۸ ارائه گردیده است که نشان می‌دهد هر متغیر چه درصدی از تغییرات متغیر وابسته را تبیین می‌نماید. به عنوان مثال، متغیر روز هفته به میزان ۱/۱ درصد و متغیر سرعت به میزان ۲/۳ درصد در احتمال وقوع هر یک از انواع خطای ثبت اطلاعات توسط دوربین‌های ترافیکی مورد مطالعه نقش دارند.

همان‌طور که مشاهده می‌شود، میزان دقت حاصل از ماتریس انطباقی مدل کلاس‌بندی k نزدیک‌ترین همسایگی برای داده‌های آزمایشی، برابر با ۶۸ درصد می‌باشد. این نتایج نشان می‌دهد که دقت این کلاس‌بند نسبتاً بالاست و متغیرهای ورودی مدل شامل شماره دوربین، ساعت حرکت خودرو، سرعت حرکت خودرو، روز هفته و خط عبور بر میزان خطای کل در ثبت اطلاعات خودرو تأثیرگذارند.

ب) رگرسیون لجستیک: نتایج حاصل از رگرسیون لجستیک در جدول شماره ۸ گزارش گردیده است. همان‌طور که از نتایج جدول مشاهده می‌شود میزان دقت (۶۰ درصد) و خطا (۲۴ درصد) در داده‌های آموزشی و داده‌های آزمایشی نزدیک هم هستند، که این یکی از معیارهای مناسب مدل است. با این وجود

جدول ۸. نتایج رگرسیون لجستیک برای بررسی میزان تأثیر متغیرهای مستقل بر خطای کل (هرگونه خطا در ثبت اطلاعات)

متغیر	ضریب تشخیص	ضرایب	انحراف معیار	Z	P-Value
ثابت رگرسیون		-۱/۵۲	۰/۰۲۴	-۶۲/۹۶۰	۰/۰۰۰
شماره دوربین X ₁	R ² = 0.000	-۰/۰۸۸	۰/۰۰۶	-۱۵/۰۶۳	۰/۰۰۰
دوشنبه X ₂₁		-۰/۳۳۹	۰/۰۱۱	-۳۲/۰۷۶	۰/۰۰۰
شنبه X ₂₂		-۰/۴۰۶	۰/۰۱۰	-۳۹/۶۸۰	۰/۰۰۰
یکشنبه X ₂₃	R ² = 0.011	-۰/۱۴۰	۰/۰۱۰	-۱۴/۲۲۸	۰/۰۰۰
پنجشنبه X ₂₄		-۰/۴۱۴	۰/۰۱۰	-۴۲/۴۱۹	۰/۰۰۰
سه‌شنبه X ₂₅		-۰/۵۰۹	۰/۰۱۰	-۴۸/۶۷۶	۰/۰۰۰
چهارشنبه X ₂₆		-۰/۶۱۰	۰/۰۱۱	-۵۷/۴۰۴	۰/۰۰۰
ساعت X ₃	R ² = 0.000	۰/۰۰۸	۰/۰۰۱	۱۳/۲۶۹	۰/۰۰۰
سرعت X ₄	R ² = 0.023	۰/۰۲۷	۰/۰۰۱	۱۰۴/۱۰۸	۰/۰۰۰
خط عبور ۲ X ₅₁		-۰/۲۰۳	۰/۰۰۶	-۳۳/۹۹۸	۰/۰۰۰
خط عبور ۳ X ₅₂	R ² = 0.004	-۰/۰۶۲	۰/۰۰۹	-۶/۶۶۷	۰/۰۰۰
خط عبور ۴ X ₅₃		۰/۰۷۹	۰/۰۲۰	۳/۹۹۷	۰/۰۰۰
حجم خودرو X ₆	R ² = 0.001	-۰/۰۱۰	۰/۰۰۱	-۱۲/۸۸۷	۰/۰۰۰

Accuracy_Test = 59.49%				Accuracy_Train = 59.47%			
MSE = 0.2376				MSE = 0.2375			
واقعی		ماتریس انطباقی داده‌های آزمایشی		واقعی		ماتریس انطباقی داده‌های آموزشی	
خطا	عدم خطا	خطا	پیش‌بینی شده	خطا	عدم خطا	خطا	پیش‌بینی شده
۳۳۱۲۱	۴۵۸۵۰	خطا	پیش‌بینی شده	۷۷۴۰۸	۱۰۶۹۷۴	خطا	پیش‌بینی شده
۹۳۰۳۹	۶۱۴۵۰	عدم خطا	شده	۲۱۶۹۶۵	۱۴۳۳۹۲	عدم خطا	شده

با توجه به ضرایب رگرسیون مندرج در جدول فوق، معادله لجیت یا معادله رگرسیون لجستیک فوق به صورت رابطه شماره ۲ است.

$$E(Y = 1|X = x) = \frac{e^{-1.5-0.09x_1-0.33x_2-0.41x_2-0.14x_2-0.41x_2-0.51x_2-0.61x_2+0.01x_3+0.03x_4-0.20x_5-0.06x_5+0.08x_5-0.01x_6}}{1+e^{-1.5-0.09x_1-0.33x_2-0.41x_2-0.14x_2-0.41x_2-0.51x_2-0.61x_2+0.01x_3+0.03x_4-0.20x_5-0.06x_5+0.08x_5-0.01x_6}}$$

۵- نتیجه گیری

در این پژوهش ۴ پیشنهاد برای بهره‌گیری پژوهشگران در زیرارائه شده است. با توجه به گستردگی و داده‌ها و موقعیت‌ها و شرایط متفاوتی که دوربین‌ها در آن قرار گرفته‌اند می‌تواند در هر دوربین تاثیر میزان هر یک از خطا‌های ارائه شده متفاوت باشد، بنابراین پیشنهاد می‌شود جهت دستیابی به دقت بالاتری از خوانش هر یک دوربین‌ها در بازه زمانی معینی مورد بررسی قرار بگیرند. -متغیرهای مورد بررسی در این پژوهش متغیرهای ثابتی در نظر گرفته شده‌اند با این حال متغیرهای دیگری نیز از جمله: نوع وسیله نقلیه اعم از سبک، سنگین و موتور سیکلت نیز می‌تواند برای روی میزان خطا نقش داشته باشند. -می‌توان نقش آلودگی هوا که باعث کاهش دید دوربین می‌شود و همچنین فرسودگی یا طول عمر تجهیزات را نیز بر روی میزان خطا دوربین‌ها مورد بررسی قرار داد. -همچنین می‌توان با مقایسه میزان خطای هر یک از دوربین‌ها با یکدیگر و نیز مقایسه شرایط فیزیکی و موقعیت دوربین شرایط فیزیکی بهینه را نیز به دست آورد.

باتوجه به اهمیت سنجش عملکرد دوربین‌های پلاک خوان شهری به منظور بررسی و شناخت صحیحی از وضعیت ترافیک شهرها، پژوهشی در این مورد انجام شده است. در این پژوهش ابتدا نمونه داده‌ها مورد بررسی و برداشت قرار گرفته‌اند. داده‌های مورد استفاده در این تحقیق مربوط به اطلاعات ترافیکی ثبت شده توسط دو دوربین ترافیکی به دوربین شماره ۱۳ و آدرس مشهد - بزرگراه شهید کلاتری، درب شرقی دانشگاه فردوسی، جنوب - شرقی و دوربین شماره ۱۴، به آدرس مشهد، بزرگراه شهید کلاتری، بلوار صدا و سیما، شمال - غربی در طول یک هفته اول خردادماه ۱۳۹۹ می‌باشد که شامل اطلاعات پلاک خودرو، تخلف ترافیکی، زمان دقیق، سرعت و تعداد خودرو است. این دو دوربین در مقایسه با دیگر دوربین‌های موجود در شهر مشهد دارای بیشترین خطا بوده‌اند. خطاهای شناسایی شده در این تحقیق شامل: خطای کل دوربین‌ها، خطای ثبت پلاک، خطای ثبت حرف پلاک، خطای ثبت کد شهر، خطای ثبت تخلف، زوج یا فرد بودن پلاک بودن است. عوامل مختلفی بر بروز این خطاها تاثیر مستقیمی داشته است از جمله: شماره دوربین، روز هفته، ساعت حرکت خودرو، سرعت حرکت خودرو، خط عبوری، تراکم یا حجم خودرو می‌باشد. هر یک از عوامل در ذیل مورد بررسی قرار داده شده است.

۶- مراجع

- Anand, S. S., & Büchner, A. G. (1998). Decision support using data mining. (No Title).
- Chitarroni, H. (2002). La regresión logística.
- Gaber, T., Abdelwahab, S., Elhoseny, M., & Hassanien, A. E. (2018). Trust-based secure clustering in WSN-based intelligent transportation systems. *Computer Networks*, 146, 151-158.
- Gunay, B., Saltan, M., Ilgaz, A., & Albayrak, Y. (2015). Area-wide ANPR coverage with a small number of instrumented vehicles. In 2015 International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS), IEEE, 379-386.

- فانژی، سید فرزین و بهروزی، اسد (۱۳۹۷). طراحی مدل مکان‌یابی شمارنده های ترافیکی با هدف استخراج حجم ترافیک در شبکه راه‌ها. *فصلنامه علمی راهور*. (۲۷)، ۱۱۰-۱۳۰.
- امیری، فریدزاد و صفارزاده، محمود (۲۰۲۰). تاثیر دوربین‌های کنترل سرعت در کاهش تخلفات سرعت با استفاده از روش بیز و گروه مقایسه (مطالعه موردی محور قزوین-رشت و بم-کرمان). *فصلنامه مهندسی حمل و نقل*، ۱۱(۳)، ۶۸۳-۶۹۶.
- میربهاء، بابک، و حسینی، سیدمهداد (۱۳۹۶). ارایه مدلی برای مکان‌یابی بهینه دوربینهای کنترل سرعت در راه‌ها. *پژوهشنامه حمل و نقل*، ۱۴(۱)، ۸۰-۶۸.

- Ebrahimzadeh, M. A., Nakhaie, N., & Dabbaqi, M. (2024). Investigating the Influential Personality and Demographic Variables of the Behavior of Public Transport and Freight Drivers in the Factors of Accidents (Case Study: Karaj City). *Road*, 32(120), 375-388.
- Ebrahimzadeh, M. A., Javanshir, H., & Esfandiyari, M. (2024). Analysis and Evaluation of Smart Ramp Metering System on Urban Traffic Flow Control (Case Study: Tehran-Hakim Highway). *Road*, 32(121), 219-234.
- Mirzahosseini, H., Ebrahimzadeh, M. A., & Amini, N. (2025). Investigating the Effect of Moonlight on Accidents on Unlit Roads. *Road*, 33(122), 199-21.
- Patel, C., Shah, D., & Patel, A. (2013). Automatic number plate recognition system (anpr): A survey. *International Journal of Computer Applications*, 69(9).
- Robinson, A., & Venter, C. (2019). Validating traffic models using large-scale Automatic Number Plate Recognition (ANPR) data. *Journal of the South African Institution of Civil Engineering*, 61(3), 45-57.
- Schreieck, M., Pflügler, C., Soto Setzke, D., Wiesche, M., & Krcmar, H. (2017). Improving urban transportation: an open platform for digital mobility services. In *Digital marketplaces unleashed*, Berlin, Heidelberg: Springer Berlin Heidelberg. 479-489.
- Tan, H., Feng, G., Feng, J., Wang, W., Zhang, Y. J., & Li, F. (2013). A tensor-based method for missing traffic data completion. *Transportation Research Part C: Emerging Technologies*, 28, 15-27.
- Ebrahimzadeh, M. A., Nakhaie, N., Dabbaqi, M., & Abdi Kordani, A. (2024). Analysis and evaluation of vehicle traffic counting in inner-city Automatic Number Plate Recognition by neural network and random forest models (Case study of Mashhad-Metropolis). *Road*.

Investigating the Reduction of the Error Caused by the Counting of Vehicle Traffic in City License Plate Reader Cameras by Mathematical Models (Case Study: Mashhad City)

Mohammad Amin Ebrahimzadeh, Ph.D., Candidate, Majoring in Transportation Planning, Faculty of Civil Engineering, Arts and Architecture, Tehran Science and Research, Tehran, Iran.

Mahdiseh Dabbaqi, Ph.D., Candidate, Majoring in Transportation Planning, Faculty of Civil Engineering, Arts and Architecture, Tehran Science and Research, Tehran, Iran.

Navid Nakhaei, Ph.D., Candidate, Majoring in Transportation Planning, Faculty of Civil Engineering, Arts and Architecture, Tehran Science and Research, Tehran, Iran.

E-mail: m.ebrahimzadeh@iau.ir

Received: January 2025- Accepted: April 2025

ABSTRACT

One of the most important ITS tools are speed control cameras, and one of the most common cameras used in urban networks is the urban license plate reader (ANPR) cameras. Therefore, due to the importance of measuring the performance of urban license plate reader cameras in order to check and correctly understand the traffic situation of cities, a research has been done to identify existing errors and then provide a mathematical model to reduce camera errors. In this research, the sample data is related to traffic information recorded by two traffic cameras in Mashhad during the first week of June 2019. The errors identified in this research include: total camera error, license plate registration error, license plate letter registration error, city code registration error, violation registration error, whether the license plate is even or odd. Various factors have had a direct impact on the occurrence of these errors, including camera number, day of the week, time of the car, and speed of the car, crossing line, density or volume of the car. In this research, in order to identify and reduce existing errors, two logistic regression and k-nearest neighbor classification models have been investigated and analyzed according to the existing parameters. In the total error, the most effective parameter is the speed and the nearest neighbor K model with 68% accuracy is the best class band for the total error of vehicle information registration.

Keywords: ANPR Camera, Logistic Regression, Traffic Situation, K Model