



دانشگاه صنعتی امیرکبیر
(پلی تکنیک تهران)

باسمه تعالی

فرم پیشنهاد پروژه کارشناسی



دانشکده مهندسی کامپیوتر

تحويل پیشنهاد پروژه به دانشکده و ثبت نهایی آن در پورتال: (این قسمت توسط کارشناسان آموزش دانشکده تکمیل می شود).

تاریخ تحويل پیشنهاد پروژه به آموزش دانشکده:

تاریخ ثبت نهایی در پورتال آموزشی دانشگاه:

مشخصات دانشجو:

نام و نام خانوادگی: نیلوفر عشقی

شماره دانشجویی: ۹۷۳۱۰۴۴

رایانامه (ایمیل) دانشجو: niloofareshghi87@gmail.com

نیمسال و سال تحصیلی ثبت نام پروژه: نیمسال دوم از سال تحصیلی ۱۴۰۰-۱۴۰۱

توضیح ۱: دانشجو موظف است حداکثر دو ماه پس از ثبت نام پروژه فرم تکمیل شده پیشنهاد پروژه را، که به امضای استاد راهنمای او رسیده است، به آموزش دانشکده تحويل دهد. انجام سر وقت این مرحله نشان دهنده بخشی از رعایت زمانبندی انجام پروژه توسط دانشجو است.

توضیح ۲: آموزش دانشکده پیشنهاد پروژه دریافتی را جهت تعیین داور و انجام داوری در اختیار گروه آموزشی استاد راهنمای دانشجو قرار می دهد. گروه های آموزشی حداکثر طی دو ماه داوری را انجام داده و در صورت تصویب در گروه، پیشنهاد پروژه را جهت تصویب در دانشکده و ثبت در پورتال آموزشی دانشگاه در اختیار آموزش دانشکده قرار می دهند. دانشجویان موظفند با داور(ان) پیشنهاد پروژه خود در ارتباط بوده و نظرات آنان را، با راهنمایی استاد راهنمای خود و در مهلت مقرر گروه برای تصویب پیشنهاد پروژه، بر روی پیشنهاد پروژه خود اعمال نمایند.

توضیح ۳: مهلت درج نموده پروژه دانشجویانی که در نیمسال اول یا در تابستان سال تحصیلی پروژه را اخذ نموده اند، سی ام مهر سال تحصیلی بعد و برای دانشجویانی که در نیمسال دوم پروژه را اخذ نموده اند، سی و یکم ام فروردین سال تحصیلی بعد است.

توضیح ۴: فاصله زمانی بین ثبت نهایی پیشنهاد پروژه (تصویب شده) در پورتال آموزشی دانشگاه و دفاع از پروژه حداقل سه ماه است و امکان دفاع قبل از سپری شدن این فاصله زمانی وجود ندارد. همچنین، دفاع از پروژه کارشناسی با اعلان عمومی و با حضور مخاطبان در حضور داوران انجام خواهد شد. لازم است دانشجویان حداقل سه هفته قبل از فرارسیدن مهلت درج نموده پروژه (توضیح ۳)، پایان نامه تایپ شده خود را، که به تأیید استاد راهنما رسیده است، در اختیار آموزش دانشکده و داور(ان) پروژه قرار داده و مقدمات برگزاری جلسه دفاع را، با هماهنگی آموزش دانشکده، فراهم آورند.

توضیح ۵: لازم است دانشجویان رویه دانشگاه صنعتی امیرکبیر با عنوان «چگونگی ثبت نام، تصویب، و دفاع از پایان نامه در مقطع کارشناسی» را که با شماره AUT-PR-3210 بر روی سایت معاونت آموزشی دانشگاه قرار گرفته است مطالعه کنند.

امضای دانشجو:

تاریخ: ۱۴۰۰/۱۲/۱۲

استاد راهنمای پروژه:

نام و نام خانوادگی:		امضا:		تاریخ:	
عنوان پروژه:					
عنوان فارسی: طراحی و پیاده سازی سامانه تحت وب هوشمند به منظور شمارش جمعیت در تصویرها					
عنوان انگلیسی: Design and Implementation of a Smart Web-based System for Crowd Counting In Images					
داور(ان) پیشنهاد پروژه:					
داور اول:					
نام و نام خانوادگی:		امضا:		تاریخ:	
داور دوم:					
نام و نام خانوادگی:		امضا:		تاریخ:	
توضیح: با امضای این قسمت داور(ان) محترم تأیید می کنند که					
۱- دانشجو، با راهنمایی استاد راهنمای خود، اصلاحات مورد نظر داور(ان) را انجام داده و عنوان و محتوای پیشنهاد پروژه از نظر ایشان قابل قبول است.					
۲- دانشجو با مفاهیم پیش نیاز و مهارت های ضروری و پایه انجام این پروژه آشنایی داشته یا کسب آن برای دانشجو در طول انجام پروژه امکان پذیر است.					
۳- موارد زیر در پیشنهاد پروژه مورد توجه قرار گرفته است:					
<ul style="list-style-type: none"> • عنوان پروژه به طور کامل و دقیق موضوع پروژه را نشان می دهد و محتوای پروژه با عنوان پروژه کاملاً مطابقت دارد. • پیشنهاد پروژه شامل بخش های مقدمه، مرور پیشینه پژوهش، رویکرد پیشنهادی، روش ارزیابی، مراحل و زمان بندی انجام پروژه، امکانات لازم و لیست مراجع و منابع است. • اجزای سامانه مورد نظر پروژه در یک نمودار بلوکی نشان داده شده و ورودی ها و خروجی های آن مشخص شده اند. • تأکید پروژه بر روی مسائل عملی و علمی و مهارت های مهندسی کامپیوتر است و پروژه منجر به توسعه نرم افزار، سخت افزار یا ترکیبی از آن دو و با درجه سختی و حجم مناسب یک پروژه سه واحدی است. • پروژه بر مبنای استفاده از دروس کارشناسی تعریف شده است. • چنانچه قرار است در پروژه از ابزارها، نرم افزارها، یا محیط های آماده استفاده شود، این موارد با صراحت بیان شده و مشخص شده است چه بخش هایی و با چه مقداری تلاش سهم دانشجو است. • پروژه علاوه بر بخش مطالعاتی-نظری، حدود ۱۵۰ ساعت کار عملی لازم داشته و انجام آن حداقل ۳ ماه زمان نیاز دارد. 					
تصویب پیشنهاد پروژه:					
تصویب در گروه آموزشی:					
نام و نام خانوادگی مدیر گروه:		امضا:		تاریخ:	
تصویب در شورای آموزشی-پژوهشی دانشکده:					
نام و نام خانوادگی معاون آموزشی:		امضا:		تاریخ:	
تعریف پروژه: (دانشجو می تواند با اضافه کردن فاصله لازم بر روی فایل قابل ویرایش این سند، توضیحات خود را در هر یک از قسمت های زیر تایپ کند).					

۱- مقدمه (بیان مسئله کاربردی، ضرورت، انگیزه، اهداف، و چالش‌های انجام این پروژه):

سیستم‌های شمارش جمعیت، یک تصویر یا ویدیو را به عنوان ورودی می‌گیرد و تعداد افراد موجود در تصویر را برمی‌گرداند. سالانه در همایش‌ها، تجمعات مردمی، مترو و رویدادهای مشابه برگزارکنندگان رویداد، شهرداری‌ها و حتی دولت‌ها با چالش بررسی، کنترل و شمارش جمعیت انبوه روبرو هستند. در دوران همه‌گیری کرونا نیز شمارش تعداد افراد در مکان‌های عمومی حتی با وجود عدم تراکم حائز اهمیت شده است. مهم‌تر از همه، یک سیستم شمارش جمعیت درست می‌تواند باعث جلوگیری از خسارات جانی و حوادث مرگبار شود. از این رو مسئله‌ی شمارش جمعیت کاربردهای مهمی دارد و سیستم‌های شمارش جمعیت می‌توانند به سازمان‌های مرتبط کمک‌رسانی کنند. در مسئله‌ی شمارش جمعیت، چالش‌هایی نظیر هم‌پوشانی چهره افراد توسط افراد دیگر یا اشیاء اطراف، شرایط نوری چالش‌زا در محیط، تنوع مقیاس، توزیع غیر یکنواخت جمعیت، کژدیسی دورنما و ... وجود دارند که غلبه کردن بر برخی از آن‌ها می‌تواند بسیار پیچیده باشد.

۲- مروری بر پروژه‌ها و سامانه‌های مشابه و بیان نقاط قوتی که با انجام این پروژه حاصل می‌شود:

رویکردهای مختلف برای شمارش جمعیت عمدتاً به چهار دسته تقسیم می‌شود: مبتنی بر تشخیص^۱، مبتنی بر رگرسیون^۲، تخمین تراکم^۳ و اخیراً رویکردهای برآورد تراکم مبتنی بر شبکه‌های عصبی پیچشی^۴ [1].

سیستم‌های مبتنی بر تشخیص با به کار بردن ابزارهای نوین تشخیص اشیاء مانند YOLO کارایی مطلوبی برای تصاویر با جمعیت پراکنده دارند [2] ولی نمی‌توانند بر چالش‌هایی که در تصاویر با جمعیت متراکم وجود دارد غلبه کنند. برای رفع این مشکل‌ها، سیستم‌های جدیدی مبتنی بر رگرسیون [3] معرفی شده‌اند که برای انجام یک نگاشت مستقیم از یک قطعه تصویر به شمار جمعیت، آموزش داده شدند. این روش‌ها به طور مرسوم ابتدا ویژگی‌های سراسری یا محلی تصویر را استخراج می‌کنند و سپس با استفاده از تکنیک‌های مختلف رگرسیون، یک تابع نگاشت از ویژگی‌های استخراج شده به شمار جمعیت، آموزش داده می‌شود. اما این روش [3] نیز به تمامی چالش‌های موجود غلبه نمی‌کند و اطلاعات مهمی مانند فاصله افراد و اطلاعات فضایی را نادیده می‌گیرد.

روش‌های جدیدتری نیز ارائه شده‌اند که برای بهره‌گیری از اطلاعات مکانی، یک نگاشت غیر خطی در کنار یک رگرسیون مبتنی بر جنگل تصادفی، برای نگاشت تصویر به یک نقشه تراکم^۵، پیشنهاد می‌دهند [4]. این دسته از روش‌های تخمین تراکم تا میزان مناسبی مشکل‌های مطرح شده را برطرف می‌کنند ولی مشکل اصلی آن‌ها استفاده از روش‌های سنتی و دستی برای استخراج ویژگی‌های سطح پایین است. کارایی استفاده از این ویژگی‌ها برای یک نقشه تراکم با کیفیت، نسبتاً کم است و همین، دقت شمارش را کاهش می‌دهد.

نهایتاً با ظهور شبکه‌های عصبی پیچشی و توانایی بالای آن‌ها در نمایش ویژگی‌ها، روش‌های نوین و کارآمد بر پایه آن‌ها پیشنهاد شده‌اند. روش‌های این دسته در طول زمان با شبکه‌های عمیق‌تر و پیشرفته‌تری طراحی شده‌اند که دقت بالایی در شمارش جمعیت‌های انبوه دارند و چالش‌های مربوط به آن را به خوبی حل می‌کنند.

یکی از روش‌های شمارش جمعیت بر مبنای شبکه عصبی پیچشی که اخیراً منتشر شده است [5]، توانسته است سیستمی طراحی کند که تصویر را به صورت یکجا و بدون تقسیم‌بندی پردازش کند. این سیستم با ترکیب ویژگی‌هایی که توسط میدان‌های پذیرندگی متفاوت بدست آمده است یک شبکه عمیق را آموزش می‌دهد. به بیان ساده‌تر این روش به صورت قابل انطباق، مقیاس اطلاعات زمینه‌ای را که برای تخمین جمعیت مورد نیاز است، در یک ماتریس مربوط به ذخیره ویژگی‌های تصویر^۶ رمزنگاری^۷ می‌کند. این روش یک روش آگاه به زمینه است که چالش تنوع مقیاس و انحراف چشم انداز را به خوبی حل می‌کند. شکل ۱ به طور خلاصه معماری این روش را توضیح می‌دهد.

¹ Detection Based

² Regression Based

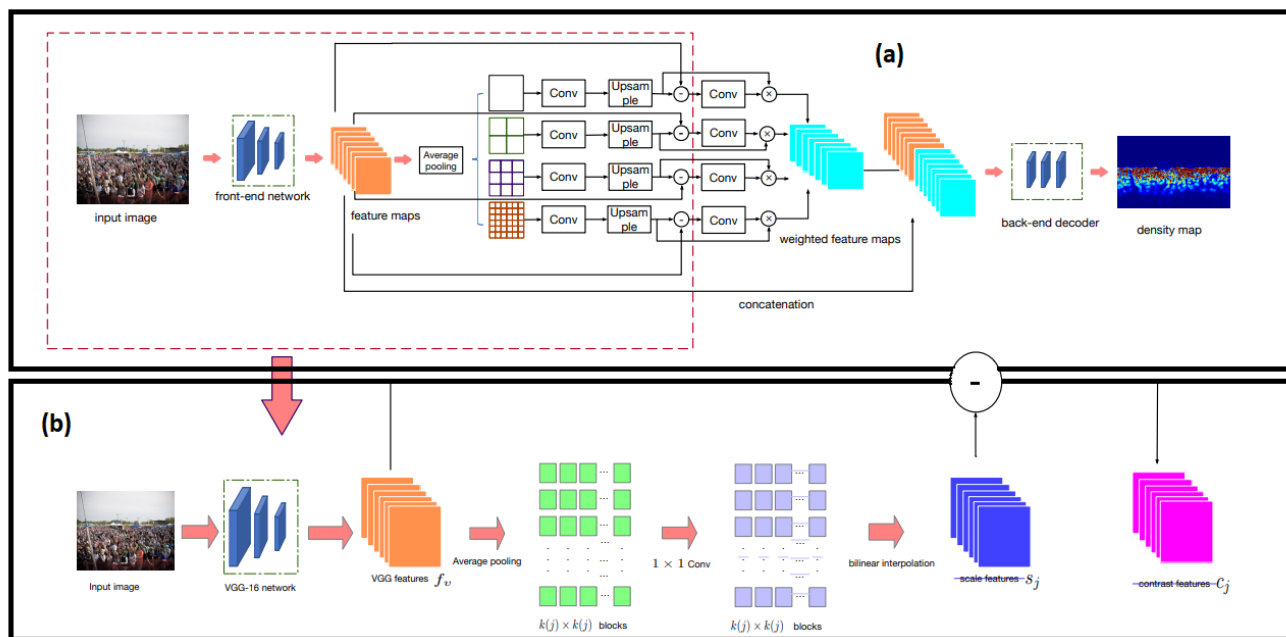
³ Density Estimation

⁴ CNN

⁵ Density Map

⁶ Feature Map

⁷ Encodes



شکل ۱: معماری یک شبکه شمارش جمعیت آگاه به زمینه/۴

در قسمت (a) شکل ۱، معماری کلی شبکه آگاه به زمینه که در کار [5] ارائه شده، آمده است. در این سیستم تصاویر با فرمت RGB به عنوان ورودی به یک شبکه عصبی پیچشی داده می‌شود که در اینجا این شبکه، ۱۰ لایه اول شبکه VGG-16 [6] است. ویژگی‌های محلی بدست آمده از این مرحله با استفاده از یک لایه رای‌گیری مبتنی بر میانگین^۸ و به دنبال آن یک لایه کانولوشن ۱ در ۱، در بلوک‌هایی با اندازه‌های متفاوت دسته‌بندی می‌شوند. سپس آنها دوباره به اندازه ویژگی‌های اصلی نمونه‌افزایی^۹ می‌شوند تا ویژگی‌های کنتراست^{۱۰} را تشکیل دهند. ویژگی‌های کنتراست در مرحله بعدی برای یادگیری وزن‌های شبکه، به منظور تولید ویژگی‌های مقیاس آگاه^{۱۱}، استفاده می‌شوند که نهایتاً برای تولید نقشه تراکم نهایی به یک شبکه پایانی رمزگشا^{۱۲} داده می‌شوند. در قسمت (b) شکل ۱، بخش اول شبکه با جزئیات بیشتری قابل مشاهده است. در این قسمت می‌بینیم که ویژگی‌های کنتراست حاصل تفریق ویژگی‌های محلی و ویژگی‌های مقیاس^{۱۳} است. این پژوهش از نظر میزان خطای تخمین جمعیت روی مجموعه داده‌هایی که در دسترس عموم است عملکرد بسیار مطلوبی داشته است و جزء ۱۰ پژوهش برتر این حوزه رتبه بندی شده است [7].

پژوهش دیگری در سال ۲۰۲۰ منتشر شده است [8] که در آن یک شبکه عصبی پیچشی بر مبنای رمزنگار-رمزگشا با استفاده از شبکه آگاه به مقیاس [5] طراحی شده است که کاستی‌های روش [5] را برطرف می‌کند. در این روش یک شبکه رمزنگار متشکل از ۱۳ لایه اول شبکه VGG16-bn^{۱۴} [6]، یک ماژول آگاه به زمینه، یک ماژول رای‌گیری هرمی مکانی^{۱۵} [9] و یک دیکدر چند مقیاسی با دو مسیر به کار برده شده است. لایه‌های شبکه عصبی VGG16-bn به تدریج اندازه نقشه ویژگی را کاهش می‌دهد و اطلاعاتی که از لحاظ معنایی سطح بالای دارند را بدست می‌آورد. ماژول‌های مقیاس آگاه نیز وظیفه حل چالش‌های مربوط به تغییرات مقیاس و کژدیسی دورنما را برعهده دارند. بخش رمزگشا نیز اطلاعات مکانی را بازیابی می‌کند تا یک نقشه چگالی نهایی با وضوح بالا تولید کند. این پژوهش علاوه بر مزیت‌های چند مقیاسی و آگاه به زمینه [5]، با استفاده از ماژول رای‌گیری هرمی مکانی، جمعیت تصاویر کم تراکم را نیز به خوبی تخمین می‌زند. در شکل ۲، معماری این روش ترسیم شده است. این روش نیز نتایج قابل توجهی داشته است و در ۱۰ روش برتر از نظر کمینه بودن میزان خطای شمارش جمعیت رتبه‌بندی شده است.

⁸ Average Pooling

⁹ Up sample

¹⁰ Contrast Features

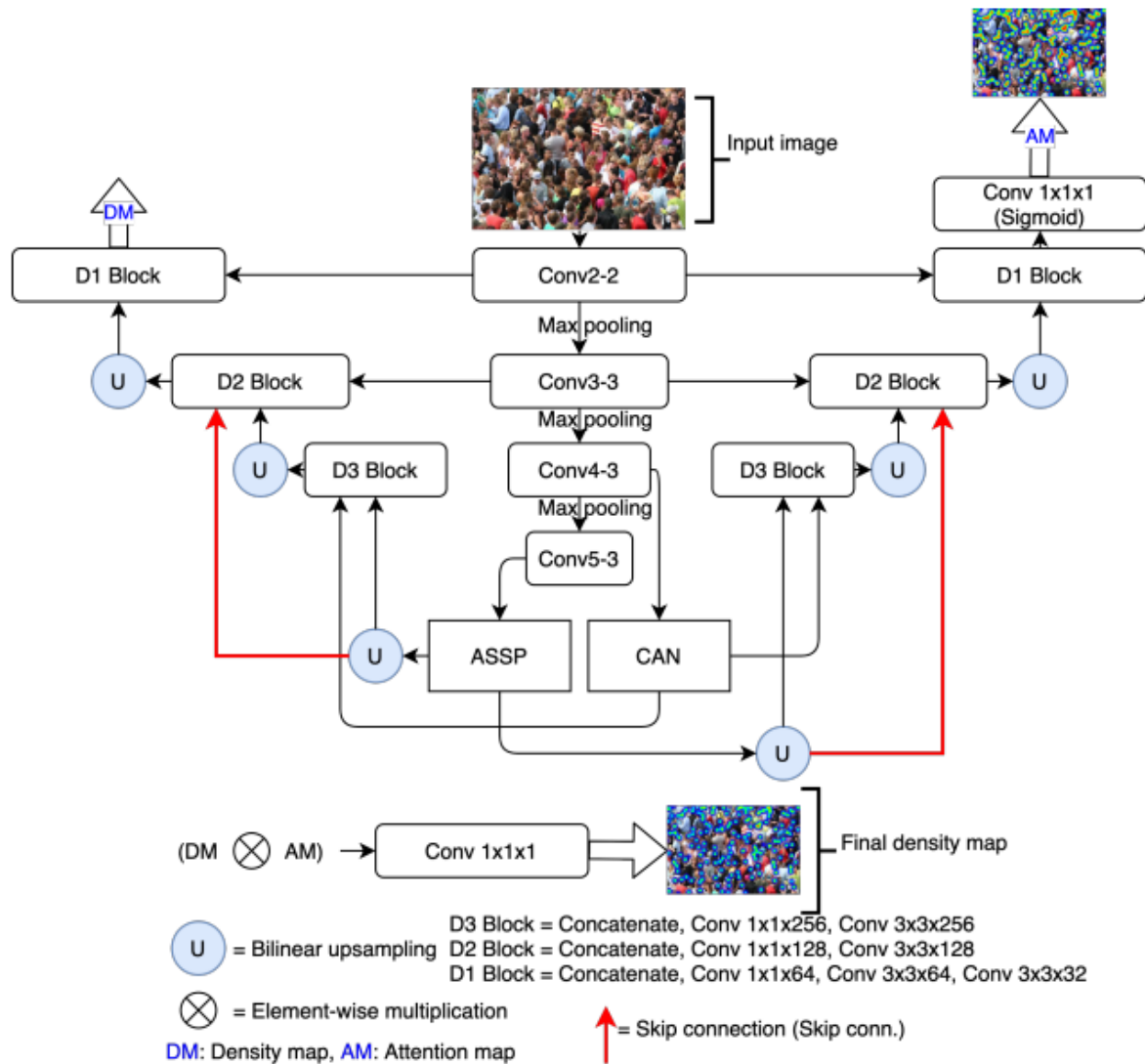
¹¹ Context-Aware

¹² Decoder

¹³ Scale Features

¹⁴ Batch normalization

¹⁵ Spatial Pyramid Pooling (SPP)



شکل ۲: معماری شبکه مبتنی بر انکدر-دیکدر [۶]

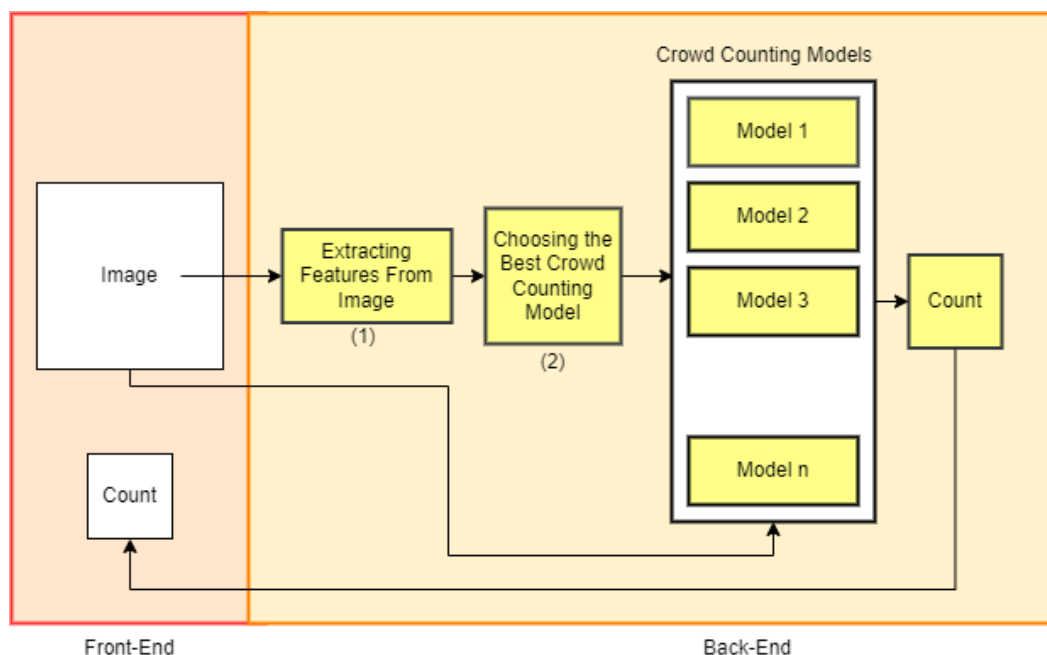
در این پروژه که در بخش بعد روش آن توضیح داده شده است، از کنار هم قرار گرفتن چند مدل شمارش جمعیت، که هر کدام در حل یکی از چالش‌های شمارش جمعیت بهترین عملکرد را دارند، استفاده می‌شود. عمل شمارش جمعیت برای هر تصویر با توجه به ویژگی‌های مختص به آن تصویر با بهترین مدل متناظر انجام می‌شود. همچنین سامانه مدنظر به صورت هوشمند تشخیص می‌دهد که کدام مدل برای کدام تصویر مناسب‌تر است و کاربر را درگیر پیچیدگی‌های انتخاب بین مدل‌های متعدد نمی‌کند. این هوشمندسازی به کاربران این اجازه را می‌دهد که تصاویر خود را به صورت انبوه به سیستم بسپارند و پاسخ مطلوبی برای هر تصویر دریافت کنند، بدون آنکه کوچک‌ترین دانشی راجع به سیستم‌های شمارش جمعیت داشته باشند. همچنین این سیستم یک سیستم تحت وب است که دسترسی آسان کاربران به آن را فراهم می‌کند.

۳- روش انجام پروژه (روش، نمودار بلوکی اجزای سامانه‌ی مورد نظر پروژه، ورودی‌ها و خروجی‌ها):

سامانه هوشمند شمارش جمعیت ما، از دو بخش سمت سرور و سمت کاربر تشکیل شده است. در سمت کاربر، یک رابط برای بارگذاری تصاویر و یک نمایشگر برای نمایش تعداد جمعیت شمرده شده وجود دارد. در سمت سرور، تصویر به یک سیستم استخراج ویژگی (بخش (۱) در شکل ۳) ورودی داده می‌شود و ویژگی‌هایی مانند میزان تراکم جمعیت، یکنواختی تراکم و ... از تصویر ورودی استخراج می‌شود. از سوی دیگر، یک سیستم هوشمند تصمیم‌گیرنده (بخش (۲) در شکل ۳) و تعدادی مدل شمارش جمعیت نیز در سمت سرور وجود دارد.

- فاز آموزش: سیستم تصمیم‌گیرنده در این سامانه می‌تواند یک شبکه عصبی بسیار ساده باشد. این سیستم با ورودی گرفتن ویژگی‌های تصویر که در سیستم استخراج ویژگی بدست آمد و با در اختیار داشتن اطلاعات عملکرد هر یک از مدل‌ها در شمارش جمعیت همان تصویر، آموزش می‌بیند. در واقع این سیستم یاد می‌گیرد که بهترین مدل شمارش جمعیت را برای هر تصویر، با توجه به ویژگی‌های استخراج شده از آن تصویر، انتخاب کند. مدل‌های شمارش جمعیت استفاده شده در این سامانه، از پیش آموزش دیده شده‌اند و در این پروژه فقط از آن‌ها استفاده می‌شود. برای آموزش سامانه و انتخاب بهترین مدل شمارش جمعیت، با استفاده از اطلاعات موجود از میزان واقعی^{۱۶} جمعیت در هر تصویر، که مجموعه داده‌های در دسترس، آن را در اختیار عموم می‌گذارند، میزان خطای شمارش سامانه محاسبه می‌شود.

- فاز کاربرد: پس از آموزش سامانه، در مرحله کاربرد، ابتدا ویژگی‌های تصویر استخراج می‌شود و سیستم تصمیم‌گیرنده (بخش (۲) در شکل ۳) متناظر با ویژگی‌های هر تصویر، بهترین مدل شمارش جمعیت را انتخاب می‌کند. سپس برای شمارش جمعیت، تصویر به مدل انتخاب شده ورودی داده می‌شود و مدل تعداد جمعیت آن را می‌شمارد. نهایتاً، تعداد جمعیت شمرده شده، خروجی داده می‌شود و تعداد جمعیت شمرده شده برای سمت کاربر فرستاده می‌شود تا در نمایشگر گفته شده، نمایش داده شود.



شکل ۳: نمودار بلوکی اجزای پروژه

همانطور که قبلاً هم اشاره شد، این سامانه تحت وب است که کاربر به عنوان ورودی تصویر را ارسال کرده و شمار جمعیت را به عنوان خروجی مشاهده می‌کند. در شکل ۳ تقسیم‌بندی قسمت‌های سمت سرور و سمت کاربر قابل مشاهده است.

۴- روش ارزیابی:

این سیستم مانند پژوهش‌های پیشین [1] از طریق دو معیار میانگین خطای مطلق و ریشه میانگین مربعات خطا ارزیابی می‌شود. این دو معیار علاوه بر مناسب بودن برای ارزیابی، به دلیل یکسان بودن با سیستم‌های دیگر کار مقایسه را ساده می‌کند. فرمول این معیارها به ترتیب در زیر آمده است.

$$MAE = \frac{1}{N} \sum_{i=1}^N |C_{I_i}^{pred} - C_{I_i}^{gt}|$$

فرمول ۱

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N |C_{I_i}^{pred} - C_{I_i}^{gt}|^2}$$

فرمول ۲

در هر دو فرمول بالا، N تعداد تصویرهایی است که در فاز تست استفاده شده‌اند، $C_{I_i}^{gt}$ و $C_{I_i}^{pred}$ به ترتیب نمایانگر نتایج تخمین و درستی مرجع هستند. به طور کلی، میانگین خطای مطلق، تعیین کننده درستی پیش‌بینی مدل است؛ در حالی که ریشه میانگین مربعات خطا، تعیین کننده قوام پیش‌بینی است.

۵- مراحل انجام و زمان‌بندی پروژه:

- انتخاب شبکه عصبی از پیش آموزش دیده شده برای استخراج نقشه ویژگی از تصویر ورودی
- انتخاب مدل‌های مناسب شمارش جمعیت از بین مدل‌های برتر
- طراحی یک سیستم هوشمند مناسب برای انتخاب مناسب‌ترین مدل شمارش بر اساس ویژگی‌های تصویر ورودی
- استخراج داده‌های آموزشی برای آموزش سیستم هوشمند و آموزش آن
- طراحی نحوه اتصال ماژول‌های بالا برای کارکرد صحیح
- پیاده سازی سیستم محاسبه خطا
- پیاده سازی سامانه سمت کاربر
- مستقر کردن سیستم روی سرور

۶- امکانات لازم (ابزارها، محیط‌ها، و نرم‌افزارهای مورد استفاده):

- Python
- Django
- PyTorch
- TensorFlow
- Numpy
- Cuda

- [1] G. Gao, J. Gao, Q. Liu, Q. Wang, and Y. Wang, "CNN-based Density Estimation and Crowd Counting: A Survey," 2020, doi: 10.48550/ARXIV.2003.12783.
- [2] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," 2016, pp. 779–788. Accessed: May 01, 2022. [Online]. Available: https://www.cv-foundation.org/openaccess/content_cvpr_2016/html/Redmon_You_Only_Look_CVPR_2016_paper.html
- [3] H. Idrees, I. Saleemi, C. Seibert, and M. Shah, "Multi-source Multi-scale Counting in Extremely Dense Crowd Images," 2013, pp. 2547–2554. Accessed: May 01, 2022. [Online]. Available: https://www.cv-foundation.org/openaccess/content_cvpr_2013/html/Idrees_Multi-source_Multi-scale_Counting_2013_CVPR_paper.html
- [4] V.-Q. Pham, T. Kozakaya, O. Yamaguchi, and R. Okada, "COUNT Forest: CO-Voting Uncertain Number of Targets Using Random Forest for Crowd Density Estimation," 2015, pp. 3253–3261. Accessed: May 01, 2022. [Online]. Available: https://openaccess.thecvf.com/content_iccv_2015/html/Pham_COUNT_Forest_CO-Voting_ICCV_2015_paper.html
- [5] W. Liu, M. Salzmann, and P. Fua, "Context-Aware Crowd Counting," 2019, pp. 5099–5108. Accessed: May 01, 2022. [Online]. Available: https://openaccess.thecvf.com/content_CVPR_2019/html/Liu_Context-Aware_Crowd_Counting_CVPR_2019_paper.html
- [6] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," Sep. 2014, doi: 10.48550/arXiv.1409.1556.
- [7] "Crowd Counting | Papers With Code." <https://paperswithcode.com/task/crowd-counting> (accessed May 01, 2022).
- [8] "Encoder-Decoder Based Convolutional Neural Networks with Multi-Scale-Aware Modules for Crowd Counting." <https://ieeexplore.ieee.org/abstract/document/9413286/> (accessed May 01, 2022).
- [9] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition," in *Computer Vision – ECCV 2014*, 2014, pp. 346–361. doi: 10.1007/978-3-319-10578-9_23.