



پروپوزال پروژه کارشناسی

Generative Adversarial Networks for persian handwritten digit

شبکه‌های مولد مقابله‌ای^۱ برای اعداد دست نویس فارسی

امیرمهدی ابوطالبی
استاد مشاور: دکتر مصطفی شمسی

خرداد ۱۴۰۲

¹Generative Adversarial Networks (GANs)

فهرست مطالب

۱	مقدمه	۲
۲	اهداف مسئله	۲
۳	روش پیشنهادی	۳
۴	پایگاه داده	۳

۱ مقدمه

در دهه‌ی اخیر، شبکه‌های مولد مقابله‌ای (GANs) به عنوان یکی از تکنیک‌های پرکاربرد در زمینه یادگیری عمیق و بینایی ماشین، توجه بسیاری را به خود جلب کرده‌اند. GANs ها قادرند تا توانایی تولید داده‌های واقع‌گرایانه را داشته باشند و از طریق دو شبکه‌ی متقابل، یک مولد^۱ [۱] و یک تمیزدهنده^۲ [۲]، بهبود بخشیدن به کیفیت تولید داده‌ها را در اختیار داشته باشند. با استفاده از شبکه‌های مولد مقابله‌ای، می‌توان داده‌های جدید و معتبری را از توزیع موردنظر تولید کرده و از آن‌ها برای موارد مختلفی مانند تولید تصاویر و یا تشخیص الگو استفاده کرد.

امروزه یکی از بزرگترین کمبودها برای ساخت مدل‌های یادگیری عمیق و یادگیری ماشین کمبود داده یا کمبود داده‌های با کیفیت است که با استفاده از GANs ها این مشکل تا حدودی رفع می‌شود. البته که مانند باقی الگوریتم‌ها و هر چیزی در جهان GANs [۳] ها نیز مشکلاتی دارند از جمله این مشکلات می‌توان به هزینه بالا برای پرورش دادن (Train) و سرعت پایین آن‌ها اشاره کرد.

در این پروژه، قصد داریم با استفاده از شبکه‌های مولد مقابله‌ای، به تشخیص ارقام دستنویس فارسی بپردازیم. در این پروژه این اعداد را به دو صورت کاملاً تصادفی و شرطی تولید می‌کنیم. که این امر باعث افزایش تعداد داده‌های این دادگاه می‌شود تا مدل‌هایی با دقت بالاتری در این زمینه تمرین داده شوند.



(آ) نمونه‌ای از داده‌های استفاده شده

۲ اهداف مسئله

- آموزش یک شبکه مولد مقابله‌ای بر روی مجموعه دادگان هدی برای تولید تصاویر ارقام دستنویس فارسی واقع‌گرایانه.
- آموزش یک شبکه تشخیص دهنده به منظور تمیزدهی بین تصاویر واقعی از ارقام دستنویس فارسی و تصاویر تولید شده توسط شبکه مولد.

^۱ Generator
^۲ Discriminator

- ارزیابی و مقایسه عملکرد مدل‌های شبکه مولد مقابله‌ای در تشخیص ارقام دستنویس فارسی با استفاده از مجموعه دادگان.

۳ روش پیشنهادی

الف) جمع‌آوری و پیش‌پردازش داده‌ها: در ابتدا، مجموعه دادگان هدی را از منبع موجود در آدرس زیر دریافت خواهیم کرد: [لینک مجموعه دادگان هدی](#). سپس، تصاویر را بررسی و پیش‌پردازش کرده و به منظور استفاده در مدل‌های GANs، آن‌ها را به فرمتی مناسب تبدیل خواهیم کرد.

ب) طراحی و آموزش مدل GANs: در این مرحله، یک مدل GANs را برای تولید تصاویر ارقام دستنویس فارسی واقع‌گرایانه طراحی خواهیم کرد. این مدل شامل دو بخش مولد و تمیزدهنده است. مولد با تولید تصاویر جدید سعی در فریب تمیزدهنده می‌کند و تمیزدهنده نیز باید بتواند تشخیص دهد که آیا تصویری که دریافت می‌کند واقعی است یا توسط مولد تولید شده است. این دو شبکه به صورت متقابل و همزمان آموزش داده می‌شوند تا بهبودی در کیفیت تولید داده‌ها حاصل شود.

ج) آموزش مدل تشخیص دهنده: پس از آموزش مدل GANs می‌توانیم از مدل تشخیص دهنده برای تمیزدهی بین تصاویر واقعی و تصاویر تولید شده توسط مولد استفاده کنیم. این مدل تشخیص دهنده را نیز با استفاده از روش‌های معمول در زمینه شبکه‌های عصبی عمیق آموزش خواهیم داد که بایستی به صورت دودویی تصمیم بگیرد که آیا تصویر دریافت شده حقیقی است یا خیر.

د) ارزیابی و مقایسه عملکرد مدل‌ها: در این مرحله، عملکرد مدل‌های GANs و تشخیص دهنده بر روی مجموعه دادگان هدی را ارزیابی خواهیم کرد. از معیارهایی مانند دقت تشخیص و میزان تطبیق با ارقام واقعی استفاده خواهیم کرده و نتایج را تحلیل خواهیم نمود.

ه) نتیجه‌گیری و پیشنهادها: در این بخش، نتایج و موفقیت‌های به‌دست آمده را خلاصه خواهیم کرد و نکاتی در خصوص پیشنهادات برای بهبود عملکرد مدل‌ها و مسیرهای آینده تحقیقاتی مشابه مطرح خواهیم نمود.

۴ پایگاه داده

پایگاه داده (HodaDataset) [۴] یک مجموعه دادگان مناسب برای تشخیص ارقام دستنویس فارسی است. این مجموعه دادگان به منظور آموزش و ارزیابی مدل‌ها و الگوریتم‌ها در زمینه تشخیص ارقام فارسی بسیار مفید است.

مجموعه دادگان هدی شامل تصاویری از ارقام دستنویس فارسی است که به صورت دستی توسط نویسندگانی که در زبان فارسی مسلط هستند، تهیه شده‌اند. تصاویر در این مجموعه دادگان شامل ارقام از ۰ تا ۹ می‌باشد که به صورت دستنویس بر روی برگه‌های کاغذی نوشته شده‌اند.

هر تصویر در مجموعه دادگان هدی ابعاد ۳۲ در ۳۲ پیکسل دارد. برای دسترسی به این تصاویر و استفاده از آن‌ها، نرم‌افزار HodaDatasetReader مورد استفاده قرار می‌گیرد. این نرم‌افزار به کمک آن، می‌توانید تصاویر را ارزیابی و به دلخواه خود عملیات پردازش و تحلیل را بر روی آن‌ها انجام دهید.

با استفاده از مجموعه دادگان هدی، می‌توان مدل‌ها و الگوریتم‌ها در زمینه تشخیص ارقام دستنویس فارسی، شبکه‌های عصبی و شبکه‌های مولد مقابله‌ای (GANs) آموزش داد و عملکرد آن‌ها را ارزیابی کرد. تعداد تصاویر در هر کلاس:

● ۱۰۰۷۰ : ۰

● ۱۰۳۳۰ : ۱

● ۹۹۲۳ : ۲

● ۱۰۳۳۴ : ۳

● ۱۰۳۳۳ : ۴

● ۱۰۱۱۰ : ۵

● ۱۰۲۵۴ : ۶

● ۱۰۳۶۳ : ۷

● ۱۰۲۶۴ : ۸

● ۱۰۳۷۱ : ۹

اطلاعات اضافی از داده:

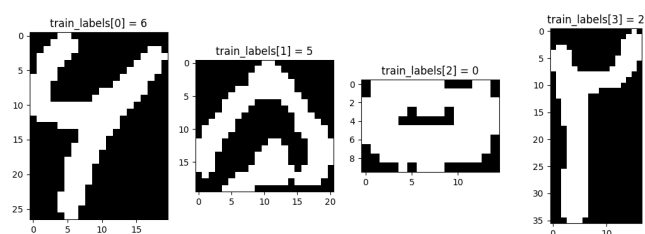
● رزولوشن هر تصویر: ۲۰۰ dpi

● تعداد نمونه‌ها: ۱۰۲,۳۵۲ samples

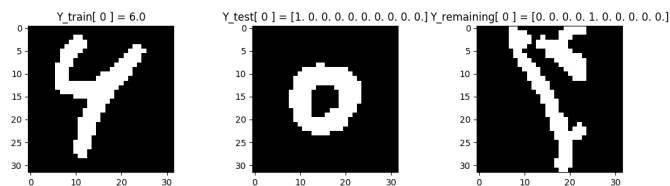
● نمونه‌های تمرینی: ۶۰,۰۰۰ samples

● نمونه‌های تستی: ۲۰,۰۰۰ samples

● داده‌های باقی‌مانده: ۲۲,۳۵۲ samples



(آ) نمونه‌ای از داده‌های استفاده شده



(آ) نمونه ای از داده های استفاده شده

مراجع

- [۱] An Introduction to Deep Generative Modeling(2021) : Lars Ruthotto, Eldad Haber [Arxiv](#)
- [۲] Discriminator-Guided Model-Based Offline Imitation Learning : Wen-jia Zhang, Haoran Xu, Haoyi Niu, Peng Cheng, Ming Li, Heming Zhang, Guyue Zhou, Xianyuan Zhan [Arxiv](#)
- [۳] Generative Adversarial Networks: Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio doi: [1406.2661](#)
- [۴] Hoda Farsi Digit Dataset GitHub: [Link](#)