Proceedings of the 2023 International Conference on Machine Learning and Automation
DOI: 10.54254/2755-2721/33/20230266

Research on the sketch recognition and automatic plotting of deep learning

Xulun Cheng

University of Virginia, 852 West Main Street, Charlottesville, Virginia

ybn8xt@virginia.edu

Abstract. Deep learning is one of the most progressive technologies around the world. With its profits of high flexibility and less economic consumption on development, more and more people are attracted to the industry and strive to improve its environment. As two of the biggest branch incorporated with such a novel and robust technology, automatic plotting, and sketch recognition have developed simultaneously. In this paper, various experiments and designs are listed to display the application of the technology. The paper indicated the possibility of such a combination through these tests and experiments. Because of the introduction of deep learning, these fields, such as facial recognition, and medicine, reduced the challenges, solved the potential risk, and paved the path for smoother improvement. Beyond the fact that deep learning still already achieves so much work, the success of in these technologies also helps its growth such as remedy the fact of limited data. It indicates its widespread usage in different areas, suggesting a confirmed future of exploiting.

Keywords: deep learning, automatic plotting, sketch recognition.

1. Introduction

As Professor Jeffery Dean stated in his work 'A Golden Decade of Deep Learning: Computing Systems & Applications', through providing a large amount of data relative to a particular purpose, the deep learning model can expand and evolve its own artificial neurons, as the Google team helped DistBelief to learn high-level concepts such as facial recognition of animals and humans through feeding it tremendous amounts of Youtube frames with collective training datas and information [1].

Currently, it is a trend that engineers prefer deep learning algorithms for developing their specialized hardware because of the two basic advantages of this algorithm. First of all, it has a high tolerance for reduced precision [1]. The deep learning model generally allows 8-bit or 4-bit integers as inference, unlike other algorithms that require 32 to 64 bits floating point numerical data [1]. In other words, the designer who applied deep learning models could fit more bits for multiplication circus on the same chip, dramatically increasing the chips 'computation speed [1]. In addition, because the chips employing deep learning could maintain more bits of low-bit computations, the chips would be more efficient and economical than normal chips [1].

With the deep learning model's advantages of high flexibility and potential, many engineers and designers are willing to keep exploring its environment, expanding its frameworks, and applying it to various fields, such as automatic plotting and sketch recognitions. According to Professor Joel Serey and their teams 'work on deep learning analysis, deep learning is applied mostly to addressing patter classification, feature learning, and feature representation for both supervised and unsupervised models [2]. Another widely used field is data management, with the mostly used function of CNN, the byproduct of deep learning algorithm, is used in 22% of the publications until 2023 February and 13% usage of DNNs [2].

2. Automatic plotting

In 2020, Professor A.Emre Kavur and their team designed a test to compare the performances of traditional semi-interactive methods and deep-learning-based automatic algorithms on determining liver volumes through computed tomography pictures [3]. The team, first, designed two groups including 6 different methods in their groups, with traditional iterative methods like water shed, fast marching, and robust statistic segmentation for chosen semi-interactive groups [3]. The team chose all the methods containing DeepMedic, NiftyNet, or U-Net templates for the deep learning group. After 12 methods were evaluated under the 20 donors 'examples [3].

Regarding accuracy and repeatability and to obtain the fusion of the individual results, Majority voting (MV) and simultaneous truth and performance level estimation (STAPLE) algorithms are applied by Professor A. Emre Kavur and their colleagues to graph the results [3].

As a result, the comparison data and individual results proved that the deep learning-based automatic plotting method is more efficient and accurate than the traditional semi-computed segmentation and volumetric measurement of livers [3]. In other words, deep learning helps refresh automatic plotting, making it more sophisticated and competitive.

Besides the result, their experiment also reinforced that, compared to the traditional plotting, several deep learning–based automatic algorithms mentioned could be applied simultaneously, implying the possibility of deep learning's expanded future [3]. Such uniqueness of the automatic plotting algorithm means that multiple algorithms could be applied in parallel, combining to form a superior performance after sorting and filtering.

As the most widely planted forage legume, alfalfa has been planted 30 million hectares worldwide. Although such expanded planting scope has proved its survivability under different situations, it is clear that one of the biggest obstacles for it to keep expanding is biomass 'labor-intensive phenotyping burden [4]. To resolve its biggest challenge, Professor Zhou Tang and their colleagues strived to apply UAV-based images with automatic plotting technology to record and analyze the potential improvement [4]. During the experiment, the team designed two groups of alfalfa fields, with the first one of 808 plots as the experiment group while and the second of 1025 plots to certify the prediction as the comparing group [4]. Such a plan was also designed with different biomass numbers to realize the labor-intensive phenotyping burden [4]. The first group experienced three harvests with biomass through UAV image automatic plotting segmentation [4]. It was recorded in three different times, May, July, and September, while the comparing group was harvested once and only recorded in September [4]. Then, the group mainly calculated four factors from the UAV image: Normalized Difference Red Edge Index (NDREI), vegetative area, plant height, and Normalized Green–Red Difference Index. As a result, the test group proves a relatively high accordance with the biomass in the second field, with around suggested 50–70% of biomass variation in the second field [4]. In other words, UAV-based automatic plotting phenotyping could enable to explicitly innovate biomass selection's efficiency under alfalfa's breeding stage.

The essential but laborious task of measuring diameters at breast height (DBH) in the traditional forest industry attracted experts to develop alternatives at the foundation of remote sensing technologies [5]. As the technology boomed in the 21st century, because of the profits behind forest surveying, The economy and high precision of structure from motion (SfM) photogrammetry attracted researchers from various fields although SfM has its shortage of unreasonable price of the light detection and ranging (LiDAR) methods [5]. Professor Gao Qiang and their team, with the technology of SfM, explored an automatic DBH measurement method by proposing an advanced technique for image acquisition and developing an automatic DBH estimation pipeline [5]. Firstly, such a novel technique focused on image obtainment could reduce the amount of resources required for effective accuracy under the DBH measurement [5]. In addition, the measurement of the DBH pipeline, which is based on RANSAC and cylinder fitting with the least median of squares, has the advantage of impressive estimation speed and high accuracy as strong as the methods based on LiDAR [5]. Moreover, the team also designed a graphical interface software Auto-DBH integrating with system introduced above in order to apply SfM on forest survey [5]. In the test, for the purpose of verifying the performance of the automatic DBH measurement, four plots are sampled with different species [5]. The result presented that the first two plots with good roundness for trees' stems has a relatively high accuracy, with its root mean squared error (RMSE) of 1.41 cm and 1.118 cm and mean relative error (MRE) of 4.78% and 5.70%, respectively [5]. However, the accuracy of the left plot with damaged trunks and low roundness for its stems was reduced, with its RMSE under 3.16 cm and MRE of 10.74% [5]. The average automatic detection rate of the trees in the four plots was 91% [5]. One of the noticeable finds is that the automatic DBH estimates relatively fast, which only consumes 2s on average to obtain the estimation of the DBH of a tree, showing its higher speed than traditional direct physical measurements [5]. More importantly, the result proves the achievement of precise result with the benefits of Auto-DBH, which includes the accuracy as high as terrestrial laser scanning (TLS) in plot scale forest DBH measurement has. Furthermore, success of the test refers that SfM has an attractive potential in forest inventory [5].

The automatic drawing and graphics storage of hydrological curves are generally difficult to work in hydrology [6]. Nowadays, although the field has developed rapidly, it still lacks the technology of automatic plotting for higher efficiency on relative works. Under the design of Professor Wang Jing and their team, the novel hydrological curve drawing software employed C++ programming language, embedded database and computer graphics technology with friendly interface and convenient operation [6]. The result, with the ability of deep learning’s automatic plotting function, included a variety of hydrological curves, including stage-discharge relation curve, index and cross-section average sediment concentration relation curve, hydrograph, channel cross-section curve and multi-year curve drawing [6]. By analyzing the precision of the stage-discharge relation curve, the program shows that all the values of systematic error, random uncertainty, mark test, curve fitting test, and deviation-data test have all met the standard requirements [6]. The development of the program fills the technical blank of the automatic plotting of hydrological curves and provides technical support for the establishment of the Yellow River hydrological electronic chart library [6].

The higher the labeling efficiency of Radiopharmaceutical is, the more the radiochemical purity of the product is [7]. Along the requirement for the purity of the product, no matter what feature the experiment or methods design to achieve, the team could not avoid the separation of the different chemical substances containing the radionuclide and estimating the percentage of radioactivity associated with the declared chemical substance [7]. However, the currently most used methods do not have specific analytical separation methods in the determination [7]. electrophoresis, gas chromatography, instant thin-layer chromatography (ITLC), liquid chromatography, thin-layer chromatography, paper chromatography, and size-exclusion chromatography are all examples of focused produced number of analytical methods [7]. In general, the mostly used methods among those are thin-layer and paper chromatography [7]. In both methods, a carrier might be added to increase volume as a remedy for tiny applied quantities of the radioactive material [7]. Thus, autoradiography or collimated counters were allowed to detect radioactive areas [7]. During the process, it created the positions of the spots or areas permitting the chemical identification after the development of the chromatographic plate [7]. Beyond the method mentioned above, Instant thin-layer chromatography (ITLC), another method commonly used, is designed for determining the labeling efficiency of radiopharmaceutical [7]. Professor Bhatwadekar, Nikhil, and their team used specific cellulose-backed silica gel chromatography strips as solid phase [7]. Such a specific design offered a variety of advantages such as easy usage, high efficiency, and less restricted incorporation into a routine quality control program [7]. Integration using an automatic-plotting instrument or a digital counter measured radioactivity [7]. The ratios of the radioactive concentration of the chemical substances is given by the ratios of the areas under the peaks [7].

DOI: 10.54254/2755-2721/33/20230266

3. Sketch recognition

By combining the traditional compound dictionary learning and deep learning, the image recognition function could achieve a new competitive level compared to other currently commonly used methods of deep learning. Professor Hao Tang and their team have proposed a novel type of deep dictionary learning and coding network (DDLCN) for image recognition without the restriction of limited data and they designed an experiment comparing it with traditional deep learning machines on image recognition [8]. To be more specific, this model employed chain-like logic, assembling multiple dictionary “atoms” under the restriction of the locality in the deep coding layer [8]. Therefore, in this system, the deep coding layer could be represented by the second one, which is designed to study the precise the fully detailed components from the components from the first layer [8]. As a result, the team created a comprehensive and discriminative low-level representation of the dictionary atoms [8]. After comparing its function with a traditional deep learning dictionary, DDLCN showed competitive data, even under the circumstance that the data is limited [8]. However, the model still needs to be improved and the only advantage currently is its well adaption of limited data [8]. With a broader range of data introduced, DDLCN has not shown enough profits for further application [8].

Beyond facial recognition technology, sketch recognition with deep learning is now introduced to its branch, makeup detection. Compared to facial recognition, makeup detection is more challenging and demanding, for its function in the field such as social interaction; cosmetology and virtual cosmetics recommendation systems; as well as security and access control; [9].

Under three different models, supervised, semi-supervised, and unsupervised models, Professor Theiab Alzahrani and their team’s test on makeup detections used labeled and unlabelled data, incorporating with the robust function of transfer learning strategy, collected and analyzed the efficiency of deep learning models for makeup detection [9]. In the supervised model, the algorithm, the VGG16 CNN, pre-trained with abundant datasets and incorporated with labeled data [9]. On the other hand, the unsupervised models, self-learning and convolutional auto-encoder, were only fed with unlabelled data [9]. Through the combination of supervised and unsupervised models, the team created the semi-supervised learning model [9]. Finally, their team’s result demonstrated that compared to supervised and unsupervised mechanisms, semi-supervised mechanisms, with the advantages of both two other models, displayed a more classification on makeup, for its 88.33% accuracy, and area under 95.15% of receiver operating characteristic curve [9]. Besides, their test shows that limited data could also affect the result of the algorithm, showing the main source, and data’s important role in developing robust deep learning.

For years, the large modality gap between a photo and a sketch within the photo-sketch matching problem is face-photo sketch recognition’s one of the biggest challenges [10]. However, Professor Seho Bae’s team has provided an advanced approach that could effectively resolve the problem by applying an intermediate latent space between the two modalities within the problem of the gap [10]. To obtain a stable homogeneous latent space between a photo and sketch for applying matching’s maximum efforts, they employed the collaborative synthesis network and equipped the latent space with rich representation power with the direction from photo to sketch and sketch and back to the photo [10]. More specifically, StyleGAN architectures like StyleGAN and StyleGAN2 are employed to gain rich representation power [10]. Therefore, such latent space with rich representation power would have the ability to conduct more accurate matching for the reason of effect alignment of distributions of the two modalities in the latent space [10]. They introduced a three-step training theme to resolve another challenge of limited paired photo or sketch samples for training [10]. As a result, their extensive evaluation of the public composite face sketch database displayed the fact that a well-trained deep learning have the possibility to gain a superior performance under the team’s approach compared to the current most widely used strategies [10]. More importantly, their success of proposing and realizing the methodology has confirmed the potential of its employment in matching other modality pairs [10].

Free-hand sketches’s effective representations are limited to learning and become a long-lasting challenge given its signal parity and the absurd-level of sketch’s abstraction [11]. The current

DOI: 10.54254/2755-2721/33/20230266

techniques exploited too much effort on employing the static nature of sketches with majorly two networks and the temporal sequential property: convolutional neural networks (CNNs) and recurrent neural networks (RNNs). However, their concentration leads to the ignorance of such an existing issue that they could not consume more effort on it [11]. According to Professor Peng Xu's team, their new work proposed a innovative achievement of sketches in the field of multiple sparsely connected graphs [11]. They designed a novel graph neural network (GNN), which is included in the multigraph transformer (MGT). In general, GNN is designed to study explicit examples of sketches from various datasets with the ability to capture comprehensive and but detailed geometric stroke structures and instant information at the same time [11]. The team, after abundant comparative experiment on GNN and MGT, demonstrated and analyzed the new model's availability and achievement [11]. In particular, around 414000 sketches from Google QuickDraw are applied to MGT [11]. As a result, the first work proposed represented sketches as graphs and applied GNNs for sketch recognition, and it has two noticeable achievements: firstly, it achieves a relatively small recognition gap that it has two percent lower than the CNN-based performance upper bound, inferring its high speed than the CNN competitors; Secondly, the method significantly outcompetes all RNN-based models [11].

Currently, most sketch recognition algorithms experts commonly examined focused on the views of the sketch as a whole instead of the different granularities during the sketching [12]. However, sketches' different granularities suggest their different roles and levels of semantic information in sketch recognition [12]. In other words, granularities' change could be referred from stroke sequences of sketches. To be more specific, as the coarser-grained contour gradually changes, it would finally evolve to a finer-grained object. To reinforce its importance and solve the potential challenge, "sketch-transfer-net" was introduced by Professor Peng Zhao's team, which is the method of transfer-deep-learning-based sketch recognition [12]. The novel precise strategy generated from sketch-transfer-net is an efficient too to adjust neural networks' various layers through different granular sketches [12]. During the process, one of the biggest challenges, a limited amount of data, is resolved by applying datasets in "ImageNet" with deep learning's robust self-training ability [12]. Such a series of extensive and provisional experiments displays that the sketch-transfer-net Peng's team designed can provide enough detained information of different granular sketches and based on the result, advance sketch recognition's conduct [12]. Additionally, this innovative accurate strategy is a potential solution to reduce the impact of transfer learning and provide a potential solution to support CNNs to be well-trained under the restriction of the current narrow database [12].

4. Conclusion

As the technology bloomed in this century, the byproduct, deep learning algorithm, has received enough resources to develop a completed environment for application. Its two biggest features, high flexibility and lower economic consumption compared to traditional machine models, helped it attract more and more people to participate in its improvement.

Nowadays, deep learning has been developed to master enough ability to cooperate with other fields to improve itself and the fields it is working with. More specifically, deep learning has more robust self-learning ability and higher efficiency than the ones traditional algorithms have, which help it develop functions such as automatic plotting and sketch recognition. In these fields, the article has listed the great achievement in recent years that deep learning has realized in these scopes. In automatic plotting, deep learning itself could achieve more laborious work in the same amount of time as the state-of-art. Beyond that, many intelligent engineers build stronger and more comprehensive systems that could meet more needs than before by combing deep learning with traditional applications, which not only make the road of the field broader but also keep the essence of the past. In sketch recognition, deep learning is one the most useful tool to help engineers and users pinpoint the target and collect the ideal data. More importantly, because of the strong connection of sketch recognition with its branches, the success of deep learning displays the potential of future employment into these scopes, helping introduce its powerful features into more demanded fields. In a word, both these fields, as two most widely applied of deep learning, have grown under the effects of such an intellectual model. Therefore, the prove of deep learning suggests its gigantic potential on other fields, even the ones that humans have not explored enough.

In the future, it is obvious that deep learning will play an important role on different systems, not only the programming system, but also the life system around the world. Deep learning models could then fully glows in all fields, starting from automatic plotting and sketch recognition.

References

[1] Dean J 2022 Daedalus 151 58–74
[2] Serey J, Alfaro M, Fuertes G, Vargas M, Durán C, Ternero R, Rivera R and Sabattin J 2023 Symmetry (20738994) 15 535
[3] Kavur A E, Gezer N S, Barış M, Şahin Y, Özkan S, Baydar B, Yüksel U, Kılıkçıer Ç, Olut Ş, Akar G B, Ünal G, Dicle O and Selver M A 2020 11–21
[4] Tang Z 1, Atit 1 P, Chen C J 1, Hu Y 1, Samuel 1 R, Medina C A 2, Sen L 2, Zhang Z 1, Long-Xi Y 2 1 W S U, Soil Sciences P and Testing Research P 2021
[5] Gao Q and Kan J 2022 Remote Sensing 14 2064–2064
[6] Jing W, Chengkun Y, Zhao Y and Weihao W. Yellow River 40 6–9
[7] Bhatwadekar N, Jain J K, Jain S K and Dubey R 2017 GMP International Journal of Pharmacy & Life Sciences 8 56–56
[8] Tang H, Liu H, Xiao W, Sebe N 2021 IEEE Transactions on Neural Networks and Learning Systems 32 2129–41
[9] Alzahrani T 1, Al-Bander B 2, Al-Nuaimy W 1 1 D of E E, Electronics U of L, pstalzah@liverpool. ac.uk 2 Department of Computer Engineering U of D 2021 497
[10] Bae S, Din N U, Park H, Yi J 2022 Sensors 22 7299
[11] Xu P, Joshi C K, Bresson X and 2022 IEEE Transactions on Neural Networks and Learning Systems 33 5150–61
[12] Zhao P, Liu Y, Lu Y and Xu B 2019 Multimedia Tools and Applications 78 35179–93