Ilkogretim Online - Elementary Education Online, 2021; Vol 20 (Issue 3): pp. 4740-4750
http://ilkogretim-online.org
doi: 10.17051/ilkonline.2021.03.485

Design And Development Of Mathematical Models For Machine Vision Applications

Robin Associate Professor, Department of Mathematics, Graphic Era Hill University, Dehradun Uttarakhand India.

Abstract

One important application of machine vision, a branch of computer vision, is in manufacturing, robotics, and autonomous systems. Machine vision systems are built on mathematical models, which provide them the ability to analyse and interpret visual data for automation and decision-making activities. The creation and development of mathematical models specifically suited for machine vision applications are discussed in this work. Image preprocessing, which is the first stage of model creation, includes converting raw visual data into a format that can be used for subsequent analysis. This covers operations like feature extraction, image improvement, and noise reduction. To produce the best results, a variety of mathematical methods are used, including edge identification, wavelet transforms, and filtering algorithms. We carry out an empirical study that contrasts several strategies in order to accomplish this goal. Through this research, we find that a hybrid strategy that substitutes simpler recurrent units for decoder self-attention, makes use of a deep encoder and shallow decoder architecture, and incorporates multiple head attention reseeding can increase accuracy. Through a harmonious fusion of time series, network design, and probabilistic solutions, we can achieve good outcomes by replacing computationally intensive functions with lighter substitutes and refining the structure of the autoencoder's layers. The results of this study show that, especially in the context of neural machine translation, careful selection and combining of methodologies can greatly enhance the performance of embedded vision applications. Developers can optimise their solutions for practical deployment by navigating the solution space and taking into account the trade-offs between speed and accuracy. In order to help developers improve the effectiveness and efficiency of their embedded vision systems through mathematical and algorithmic optimisations, this paper seeks to offer insightful advice.

Keywords: Encoder, Machine learning, Mathematical mode, activation function, computer vision.

I. Introduction

Using computers to translate statements from one natural language to another is known as machine translation (MT). Machine translation was first primarily dependent on rule-based interpretation and language expertise. However, because natural languages are so complicated by nature, it proved difficult to address all linguistic irregularities with manual conversion procedures. Data-driven methods for deriving linguistic information from data have proliferated as a result of the creation of massive parallel corpora, drawing considerable attention. In the widely used data-driven technique known [1] as statistical machine translation (SMT), latent components, such as phrase groupings or sentences, are directly acquired from parallel language corpora. SMT performs poorly when modelling long-distance dependencies between words due to its restrictions. The [2] introduction of neural machine translation (NMT) has completely changed the field and replaced SMT as the go-to MT technique. NMT uses advances in deep learning to offer a fresh method of machine translation. The translation process in NMT is based on neural networks, which can recognise intricate linguistic relationships and patterns. Large-scale parallel data training enables NMT models to develop more accurate and fluid translations. With the move towards neural machine translation, the field has advanced significantly, allowing for better translation quality and management of various language nuances [4].

Neural Machine Translation (NMT) [3,5] uses continuous representations rather than distinct symbolic representations, in contrast to Statistical Machine Translation (SMT), which uses the latter. NMT does not require considerable feature engineering because it represents the complete translation process using a single neural network. NMT training is end-to-end, in contrast to SMT, which is made up of separately created modules. NMT has made notable strides on a number of language pairings and has emerged as a crucial element in industrial machine translation systems.

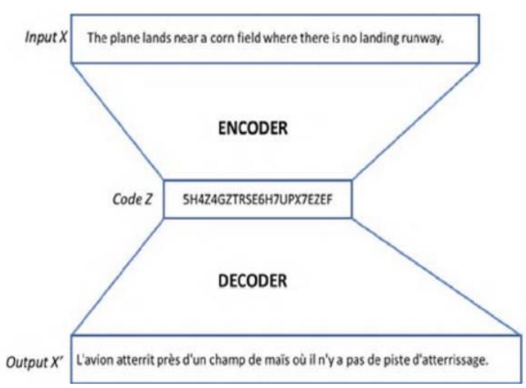

Figure 1: Encoder quadratic error

It is essential [7] to perform a thorough analysis of this strategy given the interest and variety of research areas around NMT. An overview of the core ideas and breakthroughs underlying NMT is intended in this essay. It also includes a list of freely available tools and resources for further research. By observing the progression and development of NMT. It also includes a list of freely available tools and resources for further research. The historical growth and evolution of NMT can be studied to learn important lessons that will guide future studies.

II. Related Work

The creation and improvement of mathematical models for machine vision applications has been the subject of numerous research investigations. Here, we focus on a few pertinent works in this field:

It gives a thorough analysis of the mathematical models and computational methods utilised in computer vision. It covers a wide range of topics, including 3D vision, feature extraction, and picture pre-processing. The authors explain a variety of mathematical methods and how machine vision systems can use them [8].

It focuses on leveraging the Java programming language to create machine vision algorithms. It addresses issues including object recognition, shape analysis, edge detection, and image processing. The authors present useful examples and code implementations together with the mathematical underpinnings for these techniques.

It explains mathematical [9] frameworks and methods for a range of computer vision tasks. It includes subjects like object tracking, motion estimates, segmentation, and image filtering. The authors give a thorough analysis of the mathematical ideas underlying these models and show how they work by result

The application of deep learning [10] methods in computer vision, particularly machine vision, is the main topic of this study. Convolutional neural networks (CNNs) and recurrent neural networks (RNNs), two deep learning algorithms used for tasks including image classification, object recognition, and semantic segmentation, are covered in this article's mathematical underpinnings.

It focuses on leveraging Java programming to construct machine vision algorithms. It offers detailed descriptions of various algorithms together with Java code samples. The writers describe each algorithm's use in practise as well as the mathematical ideas that underlie it. Both newcomers and seasoned professionals wishing to use Java to construct machine vision algorithms should find this book useful [11].

This review paper focuses on deep learning methods used in machine and computer vision. It gives a general overview of the mathematical underpinnings of deep learning systems like CNNs and RNNs. The benefits of deep learning in handling complicated visual data and obtaining cutting-edge performance in tasks like image categorization and object detection are covered by the authors. The limitations and potential applications of deep learning in computer vision are also highlighted in the article [12].

III. Recognition Effective Techniques

a. Statistical Machine Translation

Finding the best target sentence, indicated as y, given a source phrase, denoted as x, is how Statistical Machine Translation (SMT) [13] formulates the translation task from a source language, such as German, to a target language, such as English. By learning a probabilistic model from a parallel corpus of aligned source and target texts, SMT approaches this task.

The conditional probability of creating a target sentence given a source phrase is estimated by the probabilistic model in SMT as P(y|x). Various statistical methods and models, including n-gram language models and translation models, are used to calculate this probability.

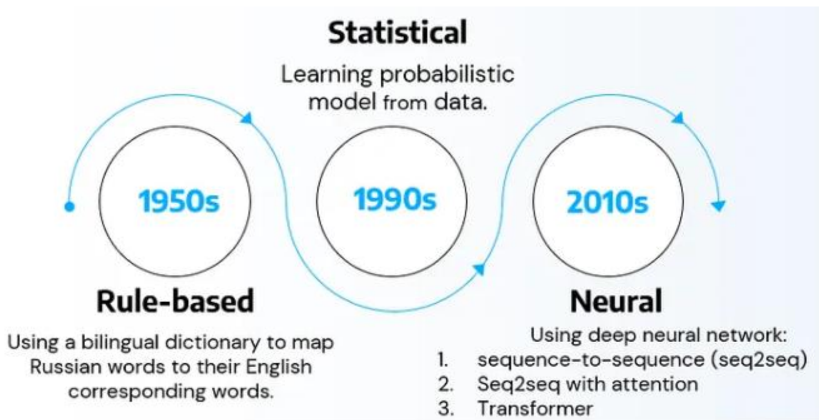

Figure 2: Machine translation overview

$$\mathrm{argmax}_y\ P(y|x)$$

Finding the best target sentence y out of all potential candidates is how Bayes' rule can be used to rephrase the translation task in statistical machine translation (SMT)

Statistical Machine Translation (SMT) uses parallel data to create a translation model called P(x|y) that attempts to capture the mapping of words and phrases between the source sentence x and the target sentence y. The word alignment or correspondence between the words in the source and target sentences is represented by the variable an in the translation model P(x, a|y), which can be further deconstructed [15].

The chance of producing strings in a particular language with monolingual input is estimated by the language model, P(y), in statistical machine translation (SMT). It is a function that gives vocabulary sequences a probability value. The probability of a language model, P(y), given a target sentence y of length n, can be calculated as follows:

$$P(y) = P(y_1, y_2, \ldots, y_n) = P(y_1) \prod_{i=2}^{n} P(y_i|y_{i:\, i-1})$$

The n-gram model in Statistical Machine Translation (SMT) can be used to approximate the probability of a word given its whole history, which addresses the inefficiencies of doing so. The n-gram model presupposes a Markov property, according to which a word's likelihood, yi, depends solely on its n-1 preceding words.

$$P(y_i|y_{i:\, i-1}) = P(y_i|y_{i-1})$$

It takes a lot of processing power to list all possible translations and determine their probability in order to discover the best translation y in statistical machine translation (SMT). SMT instead makes use of decoding [18], a heuristic search process. This approach reduces the search space by gradually eliminating hypotheses with low probabilities in order to get the optimal translation.

The fundamental SMT operating principle is described in this succinct exposition. It is crucial to keep in mind, nevertheless, that the greatest SMT systems are extremely complicated and incorporate several minute elements that have not been discussed here. Due to the requirement for considerable feature engineering and significant human work to maintain and upgrade the system over time, developing a successful SMT system can be costly and time-consuming [20].

IV. Neural Machine Translation (NMT)

Neural Machine Translation (NMT) has become a well-known topic of research in the field of employing neural networks for statistical machine translation (SMT). The sequence-to-sequence (seq2seq) model, a single end-to-end neural network architecture, is used in NMT. Two recurrent neural networks (RNNs) an encoder and a decoder make up this architecture. After processing the source sequence, the [19] encoder RNN creates an encoding vector that summarises the data from the source phrase. In contrast, the decoder RNN uses the encoding vector as input to produce the target phrase by anticipating the next word in light of the encoding vector's previous value. The seq2seq model immediately calculates the conditional probability distribution P(y|x), thereby acting as a conditional language model. Each upcoming event is predicted by the decoder RNN.

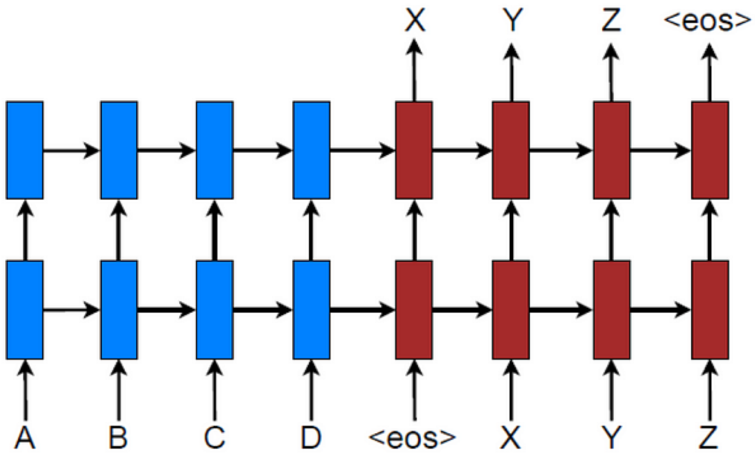

Figure 3: RNN Model as Seq2Seq for translating of machine model

The probabilistic probability distribution for the entire vocabulary of the target language will be produced by the linear layer, followed by Softmax. It will select the token with the highest probability from that probability distribution as the first word, i.e., X, and it will be utilised as the second input of the decoding [17]. The second stage of the decoder RNN will receive as inputs the second concealed state from the previous phase and the first generated word X. And until a "eos" token is generated, the same process will be repeated. The outcome of the seq2seq model will be a sequence of tokens produced by the decoder RNN.

$$\mathrm{score}(\mathrm{h}_t, \bar{\mathrm{h}}_s) = \begin{cases} \mathrm{h}_t^T \bar{\mathrm{h}}_s & \text{dot product} \\ \mathrm{h}_t^T \mathrm{W}_a \bar{\mathrm{h}}_s & \text{multiplicative} \\ \mathrm{v}_a^T \tanh(\mathrm{W}_1 \mathrm{h}_t + \mathrm{W}_2 \bar{\mathrm{h}}_s) & \text{additive} \end{cases}$$

$$\mathrm{a}_t = \frac{\exp(\mathrm{score}(\mathrm{h}_t, \bar{h}_s))}{\sum_{s'} \exp(\mathrm{score}(\mathrm{h}_t, \bar{\mathrm{h}}_{s'}))}$$

The context vector, ct, is computed using the weights from the alignment vector at. By weighing the weighted average of all the encoder hidden states, this context vector is produced. The alignment vector's weights are determined by the attention scores, which show how important each encoder hidden state was in producing the context vector. Therefore, information from the encoder hidden states with higher attention ratings dominates the context vector.

The hidden state of the decoder is then concatenated with the context vector. Similar to the basic seq2seq approach, this combined representation is utilised to choose the target language token with the highest probability. In some circumstances, the context vector from the prior phase might be added to the decoder as an extra input in addition to the decoder input.

V. Machine Vision application and Techniques Used

The study and use of computer algorithms and techniques that allow computers to derive a visual understanding from picture or video data is referred to as machine vision, also known as computer vision. It entails applying a variety of strategies to gather important data, spot patterns, and reach conclusions based on visual input. The following are some typical methods applied in machine vision applications:

Image processing: To increase the quality of photos, extract important information, or get rid of noise, image processing techniques are utilised. Before further analysis, preprocessing operations like filtering, segmentation, edge detection, and image enhancement are applied to the images.

Identifying certain items or areas of interest within an image or video is the process of object detection. To recognise and categorise objects based on their visual properties, methods including template matching, feature extraction, and machine learning algorithms (such Haar cascades, region-based CNNs) are utilised.

Image classification is the process of classifying images according to their visual content into several classes or categories. Convolutional neural networks (CNNs), for example, are machine learning methods that are frequently used to train models that can categorise images into specified classes.

Semantic Segmentation: Semantic segmentation divides an image into meaningful sections by assigning semantic labels to each pixel. For semantic segmentation tasks, deep learning approaches like fully convolutional networks (FCNs) and U-Net architectures are frequently used.

Table 1: Comparative Analysis between various machine vision applications

Technique	**Application**	**Parameters**
Image Processing	Preprocessing, noise removal	Filtering, segmentation, edge detection
Object Detection	Object recognition, tracking	Template matching, feature extraction

Image Classification	Image categorization	Convolutional neural networks (CNNs)
Semantic Segmentation	Pixel-wise labeling	Fully convolutional networks (FCNs), U-Net
Object Tracking	Object tracking in video	Correlation filters, Kalman filters
3D Reconstruction	3D object/scene modeling	Structure-from-motion (SfM), stereo vision
Deep Learning	Various machine vision tasks	Convolutional neural networks (CNNs)

The technique of finding and following a specific object or a group of objects across time in a series of frames is known as object tracking. Numerous techniques, including correlation filters, Kalman filters, and trackers based on deep learning, are used for reliable object tracking in various contexts.

Reconstruction in three dimensions (3D): From two-dimensional photographs or video data, 3D reconstruction techniques attempt to produce three-dimensional representations of objects or scenes. For 3D reconstruction, techniques like structure-from-motion (SfM), stereo vision, and depth estimates employing depth sensors (such LiDAR, depth cameras), are used.

VI. Conclusion

Computers can comprehend and interpret visual data thanks in large part to the design and implementation of mathematical models for machine vision applications. Significant progress has been made in the field of machine vision through the use of diverse approaches and algorithms, leading to state-of-the-art performance in a variety of applications. Effective language translation has been made possible by the use of statistical and probabilistic models, such as Statistical Machine Translation (SMT), which learns from massive parallel corpora. However, the creation of SMT systems is costly and time-consuming because to the substantial feature engineering and human work required. The field of machine translation has undergone a revolution with the introduction of neural machine translation (NMT). Deep neural network and sequence-to-sequence (seq2seq) architecture-based NMT models been shown to perform admirably in translation tasks. These models make use of end-to-end training, which eliminates the need for intricate feature engineering and enables a quicker and more effective translation process. Additionally, methods like attention mechanisms have been used to enhance the functionality of NMT models. During decoding, attention processes help the model produce accurate translations by focusing just on pertinent portions of the source sentence. Applications for machine vision cover a wide range of activities, such as 3D reconstruction, object identification and recognition, picture processing, and semantic segmentation. For the purpose of deriving useful information from visual input, these applications mainly rely on mathematical models and algorithms. Convolutional neural networks (CNNs) in particular have seen significant growth in deep learning, which has considerably aided the development of successful machine vision models. In comparison to conventional methods, CNNs have outperformed them in tasks involving semantic segmentation, object detection, and image classification.

References:

[1] Ba, J.L., Kiros, J.R., Hinton, G.E., 2016. Layer Normalization arXiv preprint arXiv:1607.06450.

[2] D. Bahdanau, K. Cho, Y. Bengio, Neural machine translation by jointly learning to align and translate, in: Proceedings of the 3rd International Conference on Learning Representations, 2015.

[3] M. Schuster, K.K. Paliwal, Bidirectional recurrent neural networks, IEEE Trans. Signal Process. 45 (11) (1997) 2673-2681.

[4] T. Blevins, O. Levy, L. Zettlemoyer, Deep RNNs encode soft hierarchical syntax, in Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, 2018, pp. 14-19.

[5] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in 2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016, Las Vegas, NV, USA, June 27-30, 2016, pp. 770-778.

[6] Rao, A. N., Vijayapriya, P., Kowsalya, M., & Rajest, S. S. (2020). Computer Tools for Energy Systems. In International Conference on Communication, Computing and Electronics Systems (pp. 475-484). Springer, Singapore.

[7] A. Haldorai and A. Ramu, "Security and channel noise management in cognitive radio networks," computers & Electrical Engineering, vol. 87, p. 106784, Oct. 2020. doi:10.1016/j.compeleceng.2020.106784

[8] A. Haldorai and A. Ramu, "Canonical Correlation Analysis Based Hyper Basis Feedforward Neural Network Classification for Urban Sustainability," Neural Processing Letters, Aug. 2020. doi:10.1007/s11063-020-10327-3

[9] D. Devikanniga, A. Ramu, and A. Haldorai, "Efficient Diagnosis of Liver Disease using Support Vector Machine Optimized with Crows Search Algorithm," EAI Endorsed Transactions on Energy Web, p. 164177, Jul. 2018. doi:10.4108/eai.13-7-2018.164177

[10] K. N. Durai, R. Subha, and A. Haldorai, “A Novel Method to Detect and Prevent SQLIA Using Ontology to Cloud Web Security,” Wireless Personal Communications, Mar. 2020. doi:10.1007/s11277-020-07243-z

[11] H. Anandakumar and K. Umamaheswari, “Supervised machine learning techniques in cognitive radio networks during cooperative spectrum handovers,” Cluster Computing, vol. 20, no. 2, pp. 1505-1515, Mar. 2017.

[12] H. Anandakumar and K. Umamaheswari, “A bio-inspired swarm intelligence technique for social aware cognitive radio handovers,” Computers & Electrical Engineering, vol. 71, pp. 925-937, Oct. 2018. doi:10.1016/j.compeleceng.2017.09.016

[13] Leo Willyanto Santoso, Bhopendra Singh, S. Suman Rajest, R. Regin, Karrar Hameed Kadhim (2020), “A Genetic Programming Approach to Binary Classification Problem” EAI Endorsed Transactions on Energy, Vol.8, no. 31, pp. 1-8. DOI: 10.4108/eai.13-7-2018.165523

[14] Dr. Laxmi Lidiya. S. Suman, Rajest, “Correlative Study and Analysis for Hidden Patterns in Text Analytics Unstructured Data using Supervised and Unsupervised Learning techniques” in International Journal of Cloud Computing, International Journal of Cloud Computing (IJCC), Vol. 9, No. 2/3, 2020.

[15] N. A. Jalil, H. J. Hwang, and N. M. Dawi, “Machines learning trends, perspectives and prospects in education sector,” in ACM International Conference Proceeding Series, 2019.

[16] N. A. Jalil and H. J. Hwang, “Technological-centric business intelligence: Critical success factors,” Int. J. Innov. Creat. Chang., 2019.

[17] Ravi Kumar Gupta, “Employment Security and Occupational Satisfaction in India,” Journal of Advanced Research in Dynamical & Control System, Vol. 10, Issue 10, pp. 244-249, 2018

[18] Ganguli S., Kaur G., Sarkar P., Rajest S.S. (2020) An Algorithmic Approach to System Identification in the Delta Domain Using FAdFPA Algorithm. In: Haldorai A., Ramu A., Khan S. (eds) Business Intelligence for Enterprise Internet of Things. EAI/Springer Innovations in Communication and Computing. Springer, Cham

[19] Singla M.K., Gupta J., Nijhawan P., Ganguli S., Rajest S.S. (2020) Development of an Efficient, Cheap, and Flexible IoT-Based Wind Turbine Emulator. In: Haldorai A., Ramu A., Khan S. (eds) Business Intelligence for Enterprise Internet of Things. EAI/Springer Innovations in Communication and Computing. Springer, Cham

[20] Bhopendra Singh, S. Suman Rajest, K. Praghash, Uppalapati Srilakshmi and R. Regin (2020) Nuclear structure of some even and odd nuclei using shell model calculations.

Proceedings of the 2020 2nd International Conference on Sustainable Manufacturing, Materials and Technologies. AIP Conference Proceeding