Judith Barrios

07/5/24

ITAI 1378 Computer Vision

Patricia McManus

Chihuahua vs Muffin With CNN Journal

Convolutional Neural Networks are specialized deep learning architectures created to process structured grid-like data like images. Unlike traditional neural networks that were used in the previous workshop of Chihuahua vs Muffin, which rested input data as a flat vector, meaning they convert the pixel values into a single continuous vector and represent the image in a way that is more effective to process, CNNs preserve the spatial structure of images through the many layers offered. CNNs consist of conventional layers that apply filters to be able to detect features like edges and textures, CNNs also have pooling layers that reduce dimensionality, but still retain important information, and fully connected layers for classification. The hierarchical organization and shared weights in convolution layers ensure CNNs to efficiently learn and generalize the information from the image data, making them powerful for object recognition and image classification, in contrast to traditional neural networks.

     The model’s performance during the workshop has various insights that can be drawn. The accuracy of the model for most parts is notably high, with an overall precision of around 95%, however the misclassifications reveal intriguing patterns. The model tends to confuse chihuahuas with muffins more often than the other way around, suggesting that the features that define a muffin, such as texture, shape, and color might overlap more with those of a chihuahua in certain positions or lighting areas. This occurs more frequently when the chihuahuas pose or features are identical to the muffins. For example, the chihuahua’s ears or tail can resemble the icing or crumbs of a muffin, making it a challenge for improvement on refining the model’s training dataset to better distinguish these visually similar categories.

    CNN differs drastically with traditional neural networks in performance and training time. This is due to their architecture structure that is customized for many different types of data. CNN is extremely well in spatial relationship work, examples include facial recognition or computer vision. Not every node in a CNN layer is connected to each node in the next layer, due to the convolutional layers having less parameters compared to the fully connected ones in NN. It uses convolution operations to process data by manipulating shared weights through convolutional layers, they decrease the numbers of parameters in contrast to fully connected NNS, leading to more efficiency in handling larger inputs like high-resolution images. Translating it into a faster training environment and lowers computational requirements when dealing with difficult visual patterns. Meanwhile, NNs treat each input individually, leading to less efficiency for tasks that require spatial awareness, resulting in CNN being better for tasks that handle spatial structures and outperforms in performance and reduced time.

    I was challenged to be patient and open minded with the system, because both chihuahua and muffin have similar visual features in certain poses and lighting areas, giving the model a hard time to identify between them with accuracy. Also, the difficulty of ensuring a diverse dataset that can capture wide ranges of chihuahuas and muffin images in various contexts, because it can lack diversity and have biased training, resulting in inaccurate classification. I solved these issues by continuously evaluating the model’s performance, analyzing misclassifications, and refining the dataset based on insights gained from the analysis, giving way to a more successful approach in identifying between them with more accurate results.

    Image classification that contains Convolutional Neural Networks have numerous practical applications across various fields. For example in agriculture, CNNs can analyze drone or satellite imagery to be able to monitor crop health and predict yields, helping out farmers to increase resources and provide better information on their stocks. In retail, CNNs can automate product categorization and do visual search, allowing many personalized shopping experiences for customers. They also help in security and surveillance for facial recognition and in the entertainment industry to recommend users television shows they might like based on what they have watched recently. The model has many features and abilities to help raise society’s advancement with technology and can benefit us by leveraging the power of deep learning to extract important insights from visual data.

  The ethical considerations regarding the use of the model are commonly issues of fairness, biases, and privacy. First, the issue of fairness and biases, where the concern of how systems are trained and deploy decisions that can disproportionately affect certain demographic groups. Many programmers cause bias training in data, such as underrepresenting certain skin tones or ethnicities can lead to inaccurate results in how AI system classifies images. Algorithms themselves can also cause biases if not properly designed or tested for fairness, eventually causing discrimination. For example, in facial recognition, systems have shown to perform less accurate results for dark skinned people, due to bias training. Also, privacy concerns can occur as they emerge as CNNs can inadvertently spread secret or sensitive information about individuals depicted in images, alerting questions about data protection. This can lead to distrust, security risks, including unauthorized access or breaches that could expose personal data to malicious hackers. Moreover, the ethical implications of image classification using CNNs need to ensure transparency and accountability in the systems so they are secure, justifiable, and comprehensive. These are just some of the many arising issues from society, which include potential job displacement, underscore the importance of thoughtful deployment, etc.

References:

Many of today’s new and innovative artificial intelligence (AI) applications use CNN-based deep learning technology to capture, interpret. “Convolutional Neural Networks (CNN) and Deep Learning.” *Intel*, www.intel.com/content/www/us/en/internet-of-things/computer-vision/convolutional-neural-networks.html. Accessed 5 July 2024.

Innodata Inc. “Ethical Issues in Computer Vision and Strategies for Success - INNODATA.” *Innodata Inc.*, 5 Jan. 2022, innodata.com/ethical-issues-in-computer-vision-and-strategies-for-success/.

“7 Applications of Convolutional Neural Networks - FWS.” *7 Applications of Convolutional Neural Networks - FWS*, www.flatworldsolutions.com/data-science/articles/7-applications-of-convolutional-neural-networks.php. Accessed 5 July 2024.

Zvornicanin, Written by:  Enes. “Convolutional Neural Network vs. Regular Neural Network.” *Baeldung on Computer Science*, 18 Mar. 2024, www.baeldung.com/cs/convolutional-vs-regular-nn#:~:text=This%20article%20explained%20the%20main,of%20parameters%20in%20the%20network.