CNN BASED FEATURE EXTRACTION FOR AIR POLLUTION DETECTION

R.Udaya Shanmuga

Department of Computer Science Engineering
Government college of Engineering,Tirunelveli-7

G.TamilPavai,

Department of Computer Science Engineering
Government college of Engineering,Tirunelveli-7

Abstract:
The introduction of new techniques and methods for rapidly and reliably detecting and analyzing air quality is very essential and demanding for air pollution deduction. Machine learning algorithms equipped with basic features leads to poor and slow performance due to minimal complex image characteristics representation. Now, the deep learning methodology is used as the active and effective AI tool for feature extraction. Extracting complex image features from the given images is the standard and important procedure of feature extraction which utilizes the extracted features to predict the air pollution from the given image dataset. For this purpose convolution neural network is employed in this work. Air Pollution Image Dataset (APID) was created using publicly available camera images. Fully connected layers of CNN divides the features into different categories of different classes based on similarity. The first step is to number the images into 61 groups, and the second step was to group them again into seven groups. The best results were obtained using 4096 features with an accuracy of 67 percent and 95 percent, respectively, for 61 and 7 class groups. This provides an enhanced version of feature extraction and accurate results when compared with existing methods.

Keywords- Feature extraction, CNN (convolution neural networks), SVM (support vector machine), air pollution detection

INTRODUCTION:

Recognition of air pollution from images has grown as a trending and important research topic in recent years, owing to the large number of issues caused by air pollution in metropolitan cities .Identifying which image are included in the images is one of the most difficult issues in image-based air pollution detection. This study focuses on the task of air pollution detection, assuming that given photos contain complex image features and the algorithm used to detect the air pollution. There are many techniques employed in recent days for feature extraction such as the one proposed by Adriana Romero . For remote sensing data analysis, this paper introduces the use of single-layer and deep convolutional networks. Given the high input data dimensionality and the relatively small amount of labelled data available, direct application of supervised (shallow or deep) convolutional networks to multi- and hyperspectral imagery is extremely difficult. As a result, we propose that for unsupervised learning of sparse features, we use greedy layer wise unsupervised pertaining in combination with a highly efficient algorithm. The algorithm is based on sparse representations and enforces the extracted features' population and lifetime sparsity at the same time.

Karimi [2] when properly implemented, systematic feature groupings provide a scalable, bias-free method for monitoring the degradation of laboratory and commercial systems alike. The evolution of these degradation features under various exposure conditions sheds light on the mechanisms that lead to module degradation in the field. Convolutional Neural Networks (CNN) and Support Vector Machines are the supervised algorithms used in this application (SVM). Unsupervised learning can be used to find relationships between inherent image properties as data and feature diversity grows. Feature extraction techniques aid in the detection of intrinsic geometric patterns that appear in images as a result of degradation. The extracted set of features is then subjected to principal component analysis to filter the most relevant components, which are then sent to an agglomerative hierarchical clustering algorithm.

Deep learning allows models to be better trained and identify different levels of image representation, as shown in [3]. Convolutional neural networks revolutionized image classification by learning basic shapes in the first layers and evolving to learn image features in the deeper layers, resulting in more accurate image classification. Hubel and Wiesel's work on the hierarchical representation of neurons in the visual cortex in cats inspired the concept of convolutional neural networks in 1962. Understanding the workings of the visual cortex in humans and animals was a major breakthrough in the field of computer vision. . The feature of an image is extracted in this paper using a convolution neural network and the deep learning concept. For various applications, additional classification algorithms are implemented.

The research[4] examined the peculiarities of learning data preprocessing in object data bases composed of multiple relational tables with ontology on top. Such learning data structures are unique to many novel and difficult applications. The paper proposes a new technology supported by a number of novel algorithms for ontology-centered transformation of heterogeneous possibly poor structured learning data into homogeneous informative binary feature space based on aggregation of ontology notion instances and their attribute domains and subsequent probabilistic analysis aimed at extracting more significant attributes. The proposed technology has been effectively integrated and validated across multiple case studies.

Based on these findings, a new CNN-based method for air pollution detection has been proposed. The goal of this study was to see how well a CNN model is trained previously so that it could classify images and group them into different categories using feature extraction. The next goal is evaluating the classification ability of features extracted from different fully-connected layers of CNN. In this study, features extracted from pre-trained CNN models were used to perform multi-class classification using an SVM classifier. The proposed method's results are compared to previously published results on the same image datasets in this paper.

IMPLEMENTATION:

A. Dataset:

The air pollution image dataset (APID) was used to test the algorithm developed in this paper. The dataset included images from 61 different locations and included 61 different images. Each image was taken on three different days, with six images taken from various angles and under various lighting conditions on each day. In each image, the background was kept the same, and the focus was on the object and lighting conditions. There were 1098 images in total, divided into 61 categories, in the dataset. In this study, data was divided into three folds for each lighting condition, and three-fold cross validation was used, with 12 images from two days used for training and the remaining six images used for testing. Figure 1 depicts a comparison of two images (sunny and hazy). The first three rows feature of images taken on three different days that are sunny, while the last three rows features of images taken on three different days that are hazy. Variations in shape, texture, and colour can be seen in this image for pictures of the same image taken under different lighting conditions and from different angles.

Fig 1: Division of images into several groups1

Since different images may have similar image features and physical appearance, and the training and validation images were captured on separate days with different view angles, proposed to divide images into seven different categories. "(1) foggy (2) rainy (3) snowy (4) sunny (5) windy (6) cloudy (7) hazy "

This method yielded two separate datasets, one with 61 categories and the other with seven. For both problems, two different classifiers were trained. Similar feature computation and classification. . Traditional classification methods classify images using user-generated features and classification methods (linear or non-linear classifiers). This research proposes using deep learning for feature computation and then classifying images with a linear classifier. The flow of these two methods is done by using the first layer of CNN as shown in Figure 2.

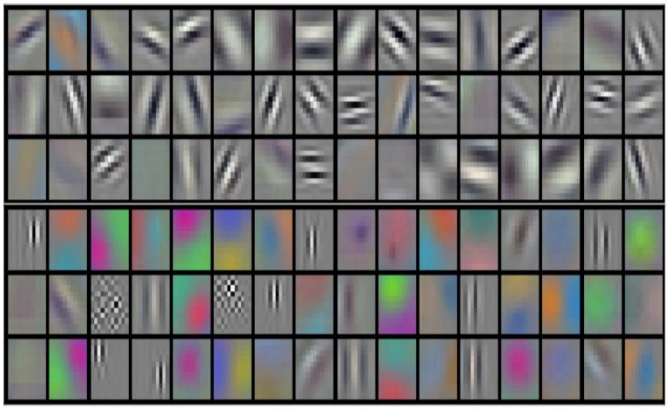

Fig 2: Filter in the first convolution layer

B. Convolution Neural Networks:

Convolutional neural networks (CNN) are the cutting-edge technology for many image recognition problems. CNNs are multi-layer neural networks with numerous convolution and pooling layers. A Convolutional neural network (CNN) is a neural network that has one or more convolutional layers and is used mainly for image processing, classification, segmentation.

- **Input layer:** The input to a CNN is mostly an image (nxmx1-gray scale image and nxmx3- colored image)

- **Convolution layer:**

Small rectangular patches make up the convolution layer (filters) smaller than the original image, with weights learned during training approaches were phases. These filters, also known as kernels, are used to extract low-level information from input images. CNN layer filters can be used to extract basic information like edges and blobs, among other things. This layer basically defines filters and computes the convolution between the defined filters and each of the images.

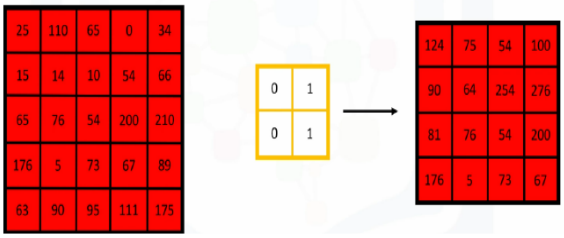

Fig. 3 convolution operation

In the same way we apply to remaining (above is for red image, then we do same for green and blue) images. We can apply more than one filter. More filters we use, we can preserve spatial dimensions better. We use convolution instead of considering flatten image as input as we will end up with a massive number of parameters that will need to be optimized and computationally expensive. E.g. this requires 25 weights if we take 5x5x1 image without convolution. We require 16 weights (n-f+1 x n-f+1) if we take 5x5x1 image with 2x2 convolution filter. By using convolution we can prevent overfitting of the model. It is worth to have ReLU activation function in convolution layer which passed only positive values and make negative values to zeros.

- **Pooling layer:**

Pooling layers are the second type of layer used by CNN. They are used to reduce the spatial size of images at each pooling layer by using some form of activation function over a rectangular window, such as maximum or average over a rectangular region. This reduces the number of parameters that must be computed, resulting in fewer computations at subsequent layers. Pooling layer objective is to reduce the spatial dimensions of the data propagating through the network. 1. Max Pooling is the most common, for each section of the image we scan and keep the highest value.

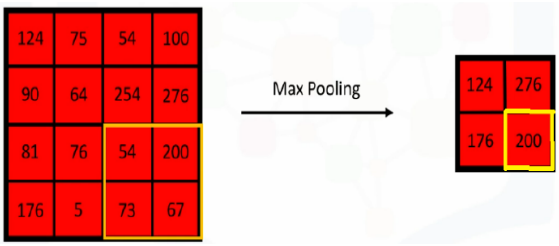

Fig. 5 Max Pooling with stride = 2

Max pooling provides spatial variance which enables the neural network to recognize objects in an image even if the object does not exactly resemble the original object. In average pooling it takes the average of area of scan. It's not advised to do Max pooling in the initial stages of the Convolutional Neural Network as the Kernels would be at the stage of extracting edges and gradients.

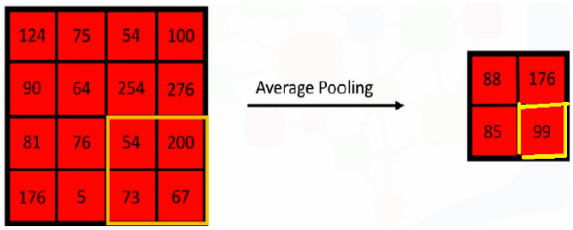

Fig.4 Average Pooling with stride = 2

- **Fully Connected Layer:**

Furthermore, a CNN architecture can have multiple fully-connected layers, similar to regular neural networks, where each layer is fully connected to all activations in the previous layer. FC represents fully-connected layers.
Here, flatten the output of the last convolutional layer and connect every node of the current layer with every other node of the next layer. This layer basically takes output of the preceding layer, whether it is a convolutional layer, ReLU or Pooling layer and outputs an n-dimensional vector, where n is number of classes pertaining to the problem. Rather than training a CNN from scratch, a pre-trained convolution neural network was used in this study. Pre-trained networks can extract features from a wide range of images. AlexNet, a network pre-trained on the ImageNet dataset, was used in this study. AlexNet has a total of 23 layers, with an input size of 227-by-227-by-3 (RGB images). Images in the PFID are 600-by-800-by-3 in size, so they were re-sampled to 227-by-227-by-3 in order to be used as an input to the network

C. Support vectors regression:

After defining the image features that could be related to PM concentration, we use a nonlinear support vector machine and a kernel to predict PM concentration. Support vector machines (SVMs) are widely used in a variety of fields, including prediction and regression. SVM can also be used to solve support vector regression problems, which are nonlinear regression estimation problems (SVR). In this study, we used SVR to forecast the PM2.5 index. The basic idea behind SVR is to use a function $\emptyset$ to map input data to a higher dimensional feature space. In the high dimensional feature space, a linear function f formulates a nonlinear relationship between input and output data. The regression function can be written as follows:

$$f(w,b) = w.\emptyset(x) + b$$

Where f (w, b) denotes the forecasting values, w and b denote the function parameter vectors, and $\emptyset$ denotes a nonlinear transformation from x to high-dimensional space. SVR's goal is to minimize function.

$$R_{reg}(f) = \frac{1}{N}\sum_{i=1}^{N} \Theta_{\varepsilon}\,(y_i w^T \emptyset(x) + b)$$

where Θ_{ε} is the ε-insensitive loss function and defined as,

$$\Theta_{\varepsilon} = \big(y, f(x)\big) = \{|f(x) - y| - \varepsilon, if |f(x) - y| \geq \varepsilon\}$$

where ε is a measure of training error, called the radius of the insensitive tube. In addition, Θ_{ε} is used to determine the optimal hyper plane in the high dimensional space and minimize the training error between the input data and the ε-insensitive loss function

The radial basis function (RBF) kernel is a popular kernel function used in regression and classification, which can handle the nonlinear relationship well, which is defined as,

$$K(x_i, y_j) = \exp(-\gamma|x_i - x_j|^2)$$

where γ is a kernel parameter. The parameters that dominate SVR are the cost constant, C, and kernel parameter, γ. We performed grid search to determine the optimal values of C and .

RESULTS AND DISCUSSIONS:

This paper presented a method for distinguishing between fast categories using a convolution neural network and linear SVM models from the APID dataset. CNN was used to automatically extract features from images rather than computing user-defined features. According to the results, the feature extracted from the FC6 fully-connected layer, combined with the Linear SVM classifier, provided the best classification results in both the 61-class classification and the 7-class classification problems as shown in table 1.

Table 1: Features extracted using CNN

S.No	Features	Accuracy
1.	Sunny	0.94
2.	Cloudy	0.72
3.	Windy	0.63
4.	Hazy	0.42
5.	Snowy	0.67
6.	Rainy	0.84
7.	Foggy	0.63

The method presented in this paper outperforms previous results on the same dataset and under similar testing conditions. For example, previous best results for a 61-class problem were reported using a combination of Pairwise Rotation Invariant Co-occurrence Local Binary Pattern (PRI-CoLBPg) features with SVM classifier, resulting in classification accuracy of 43.1 percent [14], whereas the proposed approach in with work resulted in the best accuracy of 70.13 percent, a 27 percent improvement.

Even when features from the other two layers are used, the proposed approach consistently outperforms previous approaches. (Accuracy rates of 66.39 percent and 57.2 percent, respectively). The ability of CNN to extract local and global features that are more relevant to the classification task could be one reason. Air Pollution Image Database is a difficult dataset because images were taken on three different days for each image category. Each day, photographs were taken from six different perspectives. Images were divided into seven major categories due to intra-class variations in classes:. The best previous results for 7 category classification were obtained using a combination of PRI-CoLBPg features and an SVM classifier, yielding an accuracy of 87.3 percent, where as in this study, linear SVM was used to extract features from the FC6 fully-connected layer, which resulted in a classification accuracy of 94.01 percent and a 7 percent improvement overall. Other classifiers trained with features from the FC7 and FC8 layers outperform previous results as well.

CONCLUSION:

The image dataset for this study was created using images taken using camera. This work is also important due to the widespread use of smartphones to photograph images, of the polluted surrounding environment. The method presented here can be used to recognize images and classify them into categories. The focus of future work will be on images taken during night time. Another issue is the use of learning algorithms to distinguish between indoor and outdoor air pollution, and then using algorithms to detect the level of air pollution. This will be taken into account in future work.

REFERENCES:

[1] Romero, Unsupervised Deep Feature Extraction for Remote Sensing Image Classification IEEE Transactions on Geoscience and Remote Sensing (Volume: 54, Issue: 3, March 2016)

[2] Karimi, Feature Extraction, Supervised and Unsupervised Machine Learning Classification of PV Cell Electroluminescence Images 2018 IEEE 7th World Conference on Photovoltaic Energy Conversion (WCPEC) (A Joint Conference of 45th IEEE PVSC, 28th PVSEC & 34th EU PVSEC)

[3] Manjunath Jogin. Feature Extraction using Convolution Neural Networks (CNN) and Deep Learning 2018 3rd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)

[4] J. C. Chow, J. G. Watson, J. L. Mauderly, D. L. Costa, R. E. Wyzga, S. Vedal, G. M. Hidy, S. L. Altshuler, D. Marrack, J. M. Heuss, G. T. Wolff,

[5] A. Pope, III, and D. W. Dockery, ‘‘Health effects of fine particulate air pollution: Lines that connect,’’ J. Air Waste Manage. Assoc., vol. 56, no. 10, pp. 1368–1380, Oct. 2006.

[6] Y. Xing, Y. Xu, M. Shi, and Y. Lian, ''The impact of PM2. 5 on the human respiratory system,'' J. Thoracic Disease, vol. 8, no. 1,p. E69, 2016.
[7] C. Liu, F. Tsow, Y. Zou, and N. Tao, ''Particle pollution estimation based on image analysis,'' PLoS ONE, vol. 11, no. 2, Feb. 2016, Art. no. e0145955.
[8] Z. Zhang, H. Ma, H. Fu, L. Liu, and C. Zhang, ''Outdoor air quality level inference via surveillance cameras,'' Mobile Inf. Syst., vol. 2016, pp. 1–10,Jan. 2016.
[9] X. Liu, Z. Song, E. Ngai, J. Ma, and W. Wang, ''PM2: 5 monitoring using images from smartphone Comput. Commun. Workshops (INFOCOM WKSHPS), Apr. 2015, pp. 630–635.
[10] K. Gu, J. Qiao, and X. Li, ''Highly efficient picture-based prediction of PM2.5 concentration,'' IEEE Trans. Ind. Electron., vol. 66, no. 4,pp. 3176–3184, Apr. 2019.
[11] A. Krizhevsky, I. Sutskever, and G. E. Hinton, ''ImageNet classification with deep convolutional neural networks,'' Commun. ACM, vol. 60, no. 6,pp. 84–90, May 2017.
[12] K. Simonyan and A. Zisserman, ''Very deep convolutional networks for large-scale image recognition,'' 2014, arXiv:1409.1556. [Online]. Avail- able: http://arxiv.org/abs/1409.1556
[13] K. He, X. Zhang, S. Ren, and J. Sun, ''Deep residual learning for image recognition,'' in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR),Jun. 2016, pp. 770–778.
[14] O. M. Parkhi, A. Vedaldi, and A. Zisserman, ''Deep face recogni- tion,'' in Proc. Brit. Mach. Vis. Conf., BMVA Press, 2015, p. 41, doi: 10.5244/C.29.41.
[15] H. Nam and B. Han, ''Learning multi-domain convolutional neural net- works for visual tracking,'' in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 4293–4302.
[16]T. Zheng, M. H. Bergin, S. Hu, J. Miller, and D. E. Carlson, ''Estimat-ing ground-level PM2.5 using micro-satellite images by a convolutional neural network and random forest approach,'' Atmos. Environ., vol. 230, Jun. 2020, Art. no. 117451.
[17] K. Simonyan and A. Zisserman, ''Very deep convolutional networks for large-scale image recognition,'' 2014, arXiv:1409.1556. [Online]. Avail- able: http://arxiv.org/abs/1409.1556
[18] K. He, X. Zhang, S. Ren, and J. Sun, ''Deep residual learning for image recognition,'' in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR),Jun. 2016, pp. 770–778.
[19] O. M. Parkhi, A. Vedaldi, and A. Zisserman, ''Deep face recogni- tion,'' in Proc. Brit. Mach. Vis. Conf., BMVA Press, 2015, p. 41, doi: 10.5244/C.29.41.
[20] H. Nam and B. Han, ''Learning multi-domain convolutional neural net- works for visual tracking,'' in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 4293–4302.
[21] T. Zheng, M. H. Bergin, S. Hu, J. Miller, and D. E. Carlson, ''Estimat-ing ground-level PM2.5 using micro-satellite images by a convolutional neural network and random forest approach,'' Atmos. Environ., vol. 230, Jun. 2020, Art. no. 117451.
[22] K. Y. Hong, P. O. Pinheiro, and S. Weichenthal, ''Predicting global varia-tions in outdoor PM2.5 concentrations using satellite images and deep con-volutional neural networks,'' 2019, arXiv:1906.03975. [Online]. Available:http://arxiv.org/abs/1906.03975
[23] S. Hochreiter and J. Schmidhuber, ''Long short-term memory,'' Neural Comput., vol. 9, no. 8, pp. 1735–1780, 1997.
[24] I. Sutskever, O. Vinyals, and Q. V. Le, ''Sequence to sequence learning with neural networks,'' in Proc. Adv. Neural Inf. Process. Syst., 2014,pp. 3104–3112.
[25] R. Chen, X. Wang, W. Zhang, X. Zhu, A. Li, and C. Yang, ''A hybrid CNN-LSTM model for typhoon formation forecasting,'' GeoInformatica, vol. 23, no. 3, pp. 375–396, Jul. 2019.
[26] C.-J. Huang and P.-H. Kuo, ''A deep CNN-LSTM model for particulate matter (PM2.5) forecasting in smart cities,'' Sensors, vol. 18, no. 7, p. 2220,Jul. 2018.
[27] Q. Zhang, J. C. Lam, V. O. Li, and Y. Han, ''Deep-AIR: A hybrid CNN-LSTM framework forFine-grained air pollution forecast,'' 2020,arXiv:2001.11957. [Online]
[28] H. Ozkaynak, A. D. Schatz, G. D. Thurston, R. G. Isaacs, and R. B. Husar, ''Relationships between aerosol extinction coefficients derived from air- port".

`