



MATERIAL CLASSIFICATION SYSTEM – A STUDY

Shama Holla
Department of CS
BMS College of
Engineering
Bangalore, Karnataka,
India.

Shivani Bonageri
Department of CS
BMS College of
Engineering
Bangalore, Karnataka,
India.

Shravya Shetty
Department of CS
BMS College of
Engineering
Bangalore, Karnataka,
India.

K Panimozhi
Department of CS
BMS College of
Engineering
Bangalore, Karnataka,
India.

Abstract— One of the chief challenges faced today by a visually impaired individual is to recognize a material. The inability to see and record the image structure, appearance and texture makes it difficult for them to identify an object and more importantly learn about the various kinds of materials that exist. The different material images that are being classified are glass, wood, fabric, water, stone and plastic. We have utilized the different frequency spectrum of different material images to train a machine to classify materials based on their thickness, background, number and types of covering layers and depth. The output of the image to be classified is given as a voice output, thus enabling touch-hear-comprehend learning method to be utilized for the teaching in the visually-impaired community, enabling them to be independent.

Keywords— CNN(convolutional neural network); deep learning; image classification; material recognition; insert (key words)

I. INTRODUCTION

Deriving information from images has interested several researchers for a very long time. The valuable data that can be processed from images can help solve several issues in the real world. Information about materials with different thicknesses, different backgrounds, different number and types of covering layers, and different depth distances can be extracted from images. One of the major problems is that of image classification, which is defined as predicting the class of the image. Real-world materials have geometry, rich surface texture, clutter and lighting conditions [1] which combine to make the problem particularly difficult [2],[5],[7]. Also, conversion of images existing in various formats into alternate formats to obtain the desired result, every single time, has been proven to be a tedious and redundant task. We are going to utilize different frequency spectrum of different materials to separate them through a classifier. To this end in order to address this issue, each pixel of the reconstructed image from the surface of an object under test (OUT) is taken as an input and the amplitude of that pixel in each one of 51 frequency

sample are taken as the features of that pixel, which will be used later by the trained machine for classification.[11]

This project takes an image as the input through a mobile application, processes it and classifies it into one of the six categories and gives voice output informing the category of the material it belongs to. The image given as input can be taken against different background with different lighting conditions and can be of different sizes[9]. The advancements in the field of image processing is improving day by day. Hence, it is apt that we make the most of the technology advancement to bring positive impact on lives of people who have it hard because of their disabilities. For the visually impaired, it is tough to recognize and learn about materials only by feeling it. It is also scientifically proven that hearing is one of the best methods to learn something. Keeping this in mind, we were motivated to develop a system that enables an enhanced teaching system and also empowers a visually impaired person to be slightly more independent. We are hoping the system helps build confidence and make their lives easier. Since their dependency on others is reduced while identifying materials, the chances that they are misled or cheated on would also be reduced considerably. Image processing can be a new milestone in the education system for the visually impaired and this system can be a stepping stone.

The chief objective of this project is gaining knowledge about the fast-growing technological field and understanding how it can be utilized for the betterment of the society. This project aims at developing a Classification system of materials using Deep Learning to aid a visually impaired individual to recognize material from an image. The touch-hear-comprehend feature can be used in teaching. This project has been developed with intent to utilize advancements made in technology for making the benefit of the largely populated visually impaired community. With this project we hope to introduce a new way of learning and bring about a positive change in the field of education. The proposed Material classification system aims to provide a better learning system than the traditional approach. The traditional approach does not suit every kind of person whereas the proposed system customizes the type of learning suitable for visually impaired

person whose sense of hearing and learning through touch is comparatively stronger. As the later approach customizes the learning type, it brings out a sense of confidence, independence and self-reliability.

II. PROJECT DESCRIPTION

A. Convolutional Neural Network

A convolutional neural network (CNN) is one of the important architecture of ANN (artificial neural networks) which is proposed by Yann LeCun in the year 1988. This architecture is mainly used for material recognition and image classification. Now, let us study the concept of Convolutional Neural Network in great detail. The very first step in the material recognition is accepting the input image. We humans recognize the material by looking at the image, however, computers work differently. It identifies the material with help of array of pixels. For instance, if the size of the image is 400×400 , then the size of the array will be “ $400 \times 400 \times 3$ ”. Here, the first 400 represents width, the next one represents height and 3 represents RGB channel value. RGB channel value gives the intensity of the pixel at each point which ranges from 0 to 255. Any image submitted as input to the computer will be pass through many non-linear, fully connected layer, convolutional and pooling layer to produce the required output. Computer learns the features of any image through basic things such as its boundaries and curvatures. In addition to this, many abstract concepts are constructed through series of convolutional which helps the computer to understand better [3].

Convolution layer is the first layer of CNN. The input image is first passed to this layer. Usually, the input matrix is read from top left of the image. Then, the software chooses a matrix, called as filter whose size is smaller than the previous one. This filter then produces convolution. The main task of this filter is to multiply the values of this smaller matrix with the values of the original pixel. Then sum of the obtained product is found. Similar operation is performed on the other part of the image and this is done by moving right by one unit. Once the filter covers all positions of the image, a matrix which is smaller than the input matrix is obtained. In order to identify the features of the overall image, we need network. The network contains many convolutional layers along with the pooling and non-linear layers. When the input image is passed through one of these layer, the output of this layer serves as input to the next layer. After each operation, a non-linear layer is added. This layer has activation function which adds non-linearity property. This property helps to model the response variable. After this layer, pooling layer is added. It is used to manipulate the width and height of the image and thus reduces the volume of the image. After these layers, fully connected layer is attached. The output obtained from the convolutional network is taken by this layer. After attaching this layer to the network, an N dimensional vector is obtained. Here N represents the number of classes from which the model chooses only the required class [10],[12].

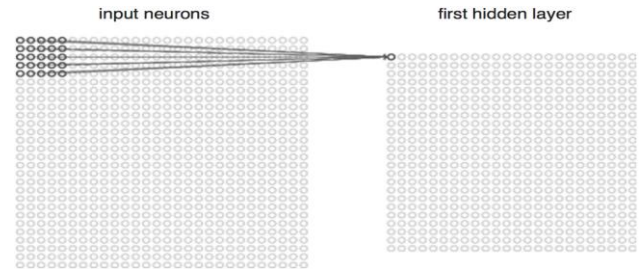


Fig. 1. Convolutional Paper

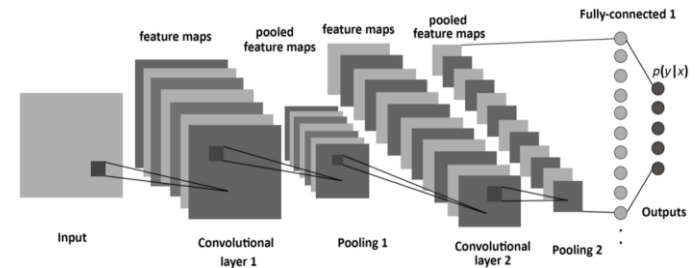


Fig. 2. CNN Architecture

B. Existing System

The existing model although has been used for several years, has a lot of drawbacks. One main disadvantage of the existing system (SVM algorithm) is that its accuracy depends on several parameters. It requires all of the key parameters to be set accurately to obtain the best result of the classification. The user is needed have to experiment with a number of different parameter values in order to achieve a satisfactory result because parameters that may result in excellent classification accuracy for one problem, may result in a poor accuracy for another. Since the number of parameters to be considered is large, a long training time is required for large datasets.

Disadvantages of existing system are: (1) Cannot properly represent the diversity of material classes with complex intra-class variability and inter-class similarity (2) Less effective for Material classification compare to human.

C. Proposed System

The new method that's proposed for this project is the convolutional neural network (CNN) [4]. They're mostly used in the analysis of visual imagery and are very efficient in image classification. CNN which can extract features without human supervision, learn and classify images based on the features. Since for classification of images, it's feature extraction is extremely important, this model serves the purpose. The features that are extracted from the images are not trained. When the network is being trained on set of images, they are learned. This increases the accuracy level of the outcomes and also is beneficial when there is a large dataset. CNN does parameter sharing and has pooling



operations enabling the system to be universal, therefore the model can run on any device.

Advantages of Proposed System are:(1)Need not figure out the features ahead of time.(2)More effective. (3)Fault tolerant.(4)Scales well

Current module is at 60% accuracy when the image is taken from the camera of the mobile device and we aim to achieve 70% accuracy. With the testing data the accuracy is currently at 80% but we are aiming to achieve up to 90% accuracy.

III. METHODOLOGY

A. Planning

Planning phase always has a very important part in the development of project . It is the very first step we need to follow to develop any successful project. It must be done in a systematic way to collect all the information and requirements of a project . Collection of data and identifying the software and hardware requirements are the two important components of the planning phase.

B. Data Collection

Images for each of the six categories- water, stone, fabric, glass, plastic and wood are collected. Each dataset has about 400 images for each of the six categories.

C. Data Pre-Processing

The raw data that has been collected is pre-processed to convert into a clean dataset. Images are binarized and the raw data is resized and normalized to remove noise, inconsistency, bad quality images and duplication. OpenCV library in python is used for this process.

D. Classification of images using Convolution Neural Network(CNN):

Deep Learning algorithms have great potential for research into the automated extraction of complex data from images. Deep Learning algorithms can develop a layered and hierarchical architecture of learning. Deep Learning allows extraction of complex abstractions through a hierarchical learning process. A key benefit of Deep Learning is that it can learn from large volume of unsupervised data.

Generally, CNN has two layers -feature extraction layer and feature map layer. In the extraction layer, the input of each neuron is connected to the previous layer receptive and feature will be extracted. After the features are extracted, positional relationship between the extracted features and will be determined. In feature map layer, each layer has several feature maps. The weight of the neurons in each feature map are equal. of The feature map uses the sigmoid function to make the shift invariance in it. Each convolution layer is followed by a computing layer which calculates the average. This model reduces the resolution. The feature extraction layer

of CNN model learns by the feature vectors which are served as training dataset. CNN algorithm is more efficient because it avoids extraction of keywords. It will classify the image based on learning from the dataset. This algorithm has an accuracy of about 55%. When a new set of images are given as the input to model, model will classify it.[6],[8]

E. Testing and evaluation of the model

About 80% of the data is dedicated to the training model while 20% is used to test the accuracy of the trained model. The data in the testing model has never been exposed to the CNN algorithm hence the features and attributes have to be identified accurately based on the learning of the system with training data. The output given is compared with the actual result in the evaluation process to check how correct the model is.

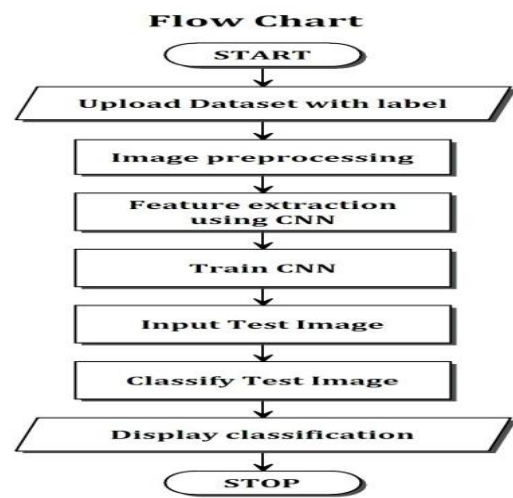


Fig. 3. Methodology Flow chart

IV. ANALYSIS

In the final phase, classification model is tested using prepared image dataset and also performance is evaluated. After developing the predictive model, it is tested using the data that was not used during the training process of the model. We can use different models for the same prediction problem which helps us to identify the right model for the given problem based on their performance. Accuracy is one of the commonly used metrics to evaluate performance of a predictive model. Co-incidence matrix or classification matrix is the primary source of performance measurements in any classification problems. Co-incidence matrix for a “two class” classification problem and the equations of the most commonly used metrics that can be calculated from the coincidence matrix are given in Fig 4 and Fig 5. The numbers that are along the principal diagonal of the matrix (in the Fig 4) represent the correct values where as the numbers along the other diagonal represent errors.



		True Class	
		Positive	Negative
Predicted Class	Positive	True Positive Count (TP)	False Positive Count (FP)
	Negative	False Negative Count (FN)	True Negative Count (TN)

Fig. 4. Confusion matrix

$$\text{True Positive Rate} = \frac{TP}{TP + FN}$$

$$\text{True Negative Rate} = \frac{TN}{TN + FP}$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

Fig. 5. Formulas used for the analysis

V. CONCLUSION

The development and implementation of material classification system is an important milestone in the education and for the daily life convenience of the visually impaired community. Many individuals facing the disability have a difficult time in learning and recognizing various objects because of their inability to visualize. As students who are studying about the technology, it's our responsibility to develop a system that can help make their lives slightly easier. Proposed system is based on the concept of convolutional neural networks and mobile app development to ensure efficiency and accuracy in the output. The proposed system put forwards a deep learning based system which will automatically categorize materials based on the trained set of data.

Hence, this system helps in building confidence, enables them to be independent and also can possibly be a new milestone in the education system for the visually impaired.

VI. REFERENCES

[1] Xue, J., Zhang, H., Dana, K., & Nishino, K. (2017). Differential angular imaging for material recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 764-773).

[2] Schwartz, G., & Nishino, K. (2019). Recognizing material properties from images. *IEEE transactions on pattern analysis and machine intelligence*.

[3] Schmidt, F., & Fleming, R. W. (2018). Identifying shape transformations from photographs of real objects. *PLoS one*, 13(8).

[4] Sticlaru, A. (2017). Material Classification using Neural Networks. arXiv preprint arXiv:1710.06854.

[5] Schwartz, G., & Nishino, K. (2016). Material recognition from local appearance in global context. arXiv preprint arXiv:1611.09394.

[6] Kalliatakis, G., Sticlaru, A., Stamatiadis, G., Ehsan, S., Leonardis, A., Gall, J., & McDonald-Maier, K. D. (2017). Material Classification in the Wild: Do Synthesized Training Data Generalise Better than Real-World Training Data?. arXiv preprint arXiv:1711.03874.

[7] Schwartz, G., & Nishino, K. (2016). Integrating Local Material Recognition with Large-Scale Perceptual Attribute Discovery. arXiv preprint arXiv:1604.01345.

[8] Bell, S., Upchurch, P., Snively, N., & Bala, K. (2015). Material recognition in the wild with the materials in context database. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3479-3487).

[9] Wang, O., Gunawardane, P., Scher, S., & Davis, J. (2009, June). Material classification using BRDF slices. In 2009 IEEE Conference on Computer Vision and Pattern Recognition (pp. 2805-2811). IEEE.

[10] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

[11] Hu, D., Bo, L., & Ren, X. (2011). Toward Robust Material Recognition for Everyday Objects. In *BMVC* (Vol. 2, p. 6)..

[12] Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 580-587).