A Convolutional Neural Network Approach to the Classification of Engineering Models

BHARADWAJ MANDA[ID]**, PRANJAL BHASKARE, AND RAMANATHAN MUTHUGANAPATHY**

Advanced Geometric Computing Laboratory, Department of Engineering Design, Indian Institute of Technology Madras, Chennai 600036, India

Corresponding author: Bharadwaj Manda (ed16d405@smail.iitm.ac.in)

ABSTRACT This paper presents a deep learning approach for the classification of Engineering (CAD) models using Convolutional Neural Networks (CNNs). Owing to the availability of large annotated datasets and also enough computational power in the form of GPUs, many deep learning-based solutions for object classification have been proposed of late, especially in the domain of images and graphical models. Nevertheless, very few solutions have been proposed for the task of functional classification of CAD models. Hence, for this research, CAD models have been collected from Engineering Shape Benchmark (ESB), National Design Repository (NDR) and augmented with newer models created using a modeling software to form a dataset - 'CADNET'. It is proposed to use a residual network architecture for CADNET, inspired by the popular ResNet. A weighted Light Field Descriptor (LFD) scheme is chosen as the method of feature extraction, and the generated images are fed as inputs to the CNN. The problem of class imbalance in the dataset is addressed using a class weights approach. Experiments have been conducted with other signatures such as geodesic distance etc. using deep networks as well as other network architectures on the CADNET. The LFD-based CNN approach using the proposed network architecture, along with gradient boosting yielded the best classification accuracy on CADNET.

INDEX TERMS Engineering/CAD models, classification, convolutional neural network, gradient boosting, light field descriptor (LFD).

I. INTRODUCTION

Classification of Engineering (CAD) models is very important for a task such as design reuse. It has been observed that designers spend a considerable amount of time in search for the right information as well as use a large percentage of existing design for a new product development [1]. Gunn [1] has observed that about 40% of the new designs could be built from an existing design and 40% from modifying an existing design. Ullman [2] has indicated that a large percentage (75% or sometimes, more than that) of design reuses existing knowledge for the new product development. Classification is also an important task for retrieval of CAD models, which in turn employed in design reuse [3]. Another area of interest is in the CAD assembly model retrieval [4], where, apart from using topology and connection informations, classification plays a key role. The interdependence between product life cycle management (PLM), material requirements planning (MRP) and CAD systems also calls for classification and search of 3D Engineering models [5].

The associate editor coordinating the review of this manuscript and approving it for publication was Huanqing Wang.

Considering the applications and the fact that we are in the digital age with many information archived digitally, the problem of automatic classification of CAD models becomes a predominant one.

An Engineering/CAD model (see Fig. 1a) has features such as holes (genus > 0), blind holes (genus = 0) and fillets which are usually absent in a graphical model (see Fig. 1b). Also, sharp edges are usually found in a CAD Model as opposed to a graphical model which more or less has smooth curvature throughout.

(a) A 3D CAD Model   (b) A 3D Graphical Model
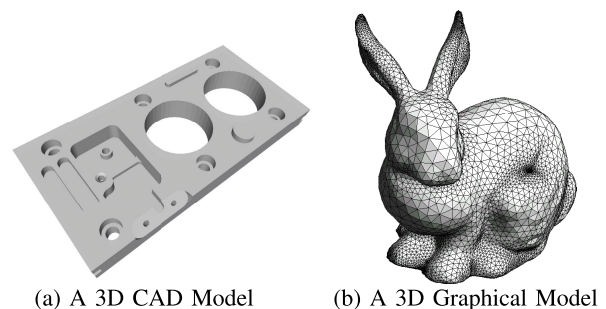
FIGURE 1. Distinction between a CAD model and a graphical model.

Traditionally, in the field of Engineering models, boundary representation (B-Rep) is the native format to store the data. To protect the proprietary design during data transfer, other formats such as mesh / tessellated representations are used. Also, with newer technologies such as additive manufacturing / 3D printing, mesh model representation in the field of Engineering / CAD is gaining popularity.

A mesh for a CAD model (see Fig. 1a) is substantially different from that of a graphical model (see Fig. 1b) in the following ways [6]:

- CAD Model tessellations are typically sparse while that of graphical models are usually dense.
- CAD Models have a lesser number of triangles in general as compared to a graphical model.

In general, the problem of classification of shapes/models has been an active area of research in several fields viz. images, graphical models, CAD models etc. for more than two decades. The problem gained prominence with the start of digital archiving. For image data, MPEG dataset [7] was perhaps one of the first ones in the domain of computer vision. In the area of graphical models, Princeton shape benchmark (PSB) was one of the earlier ones [8]. Jayanti et al. introduced Engineering shape benchmark (ESB) for CAD models [6]. The National Design Repository (NDR) [9] also contains a few hundred CAD models.

The advent of machine learning techniques and in particular, the advances made in deep learning, accelerated the research on the problem of classification. However, deep learning techniques call for a large number of labeled data with appropriate class information. Hence, labeled datasets for images with class information have grown much larger [10], and so for graphical 3D models [11]. In recent times, datasets for Engineering/CAD models have also grown through acquiring from different resources [12]. Though the dataset [12] contains a large number of models, the aim seems to be more on populating the data rather than providing classification. In the field of Engineering, it is imperative to classify data functionally. For example, a pipe and a bolt may look like a cylinder but have different functionality. This classification task also requires the people involved to have rich domain knowledge and experience. As CAD models are a derivate of the Engineering design process, many of the design data are also proprietary in nature and hence may not be put in the public domain [13]. Also, there exists only very few works on CAD model classification using deep learning. Qin et al. [13] use only deep networks (not CNN) and the work presented in [12] uses basic geometric properties such as normals and curvatures but does not take functional classification into account.

Our motivation for addressing the problem of classification of Engineering / CAD models comes from the following:

1) Most of the CAD datasets (such as [6] or [9]) have only a few hundred models.
2) Datasets having larger number of CAD models are either proprietary (not publicly available) [13] or lack classification information [12].

3) The recent advances in deep learning such as CNN have not been made use of, to the best of our knowledge.

In this work, in order to classify the CAD data functionally, we start by using the publicly available datasets of CAD Models, ESB and NDR, which also have well-annotated functional classification. Unfortunately, they have only very few models, in the order of hundreds. We then resort to creating CAD models and functionally classify them by adding to the appropriate class. A dataset termed 'CADNET' has been then prepared. A Convolutional Neural Network (CNN) approach (only deep network was used in [13]) for the classification of CAD models is then proposed. It is also crucial to come up with a network architecture for the intended application. The key contributions of the paper are as follows:

1) A dataset named 'CADNET', which is suitable for deep learning-based approaches.
2) A CNN-based deep learning approach for the classification of CAD models using a residual network structure, inspired by ResNet [14], with much lesser number of filters and thereby much reduction in the number of parameters.
3) Used the idea of class-weights, in order to alleviate the problem of imbalanced classes in CADNET.
4) Proposed the use of gradient-based boosting approaches to improve the efficiency of classification.
5) The proposed network produces better classification accuracy with much lesser training time.

The manuscript is organized as follows: Section II discusses the literature corresponding to 3D CAD models, in addition to the literature on Images, 3D Graphical Models, and an overview of the existing datasets for CAD model classification. The typical pipeline to be employed for a deep learning based classification task is described in Section III, and each step of the pipeline is elaborated in Sections IV to VI. Section VII provides the Implementation Details. The results, limitations and possible future work are elaborated in Section VIII, followed by a Conclusion (Section IX).

II. RELATED WORKS

Many works in recent times have focused on 3D graphical models and images. However, we focus more on the approaches that have been proposed for the task of classifying 3D CAD models, which are very few.

A. 3D CAD MODELS

Wu and Jen [15] proposed a neural network approach to classify 3D prismatic parts. The idea used the hypothesis that a 3D part could be modeled by the contours of its three projected views. The views were then approximated by as many rectilinear polygons. Such a representation was used as an input vector to a back-propagation neural network. A total of 36 parts were classified in this way but the classification was hierarchical and not functional. The research presented in [16] and [17] aimed at performing a classification of the CAD models based on their manufacturing process as well as

TABLE 1. Summary of existing datasets indicated the advantages and limitations of each of them. Most of the datasets are small as indicated by their respective numbers.

Reference	Total Num. of Models	Num. of classes	Avg. Num. of models per class	Comments
Wu et al. [15]	36	–	–	(-) contains only one class
Ip et al. [16]; & Ip et al. [17]	85	12	7	(+) functional classification (-) less number of models; not suitable for deep learning;
Ip et al. [16]; & Ip et al. [17]	56	4	14	(-) too few models (-) manufacturing classification;
Ip et al. [18]	100	–	–	(-) single category; (-) manufacturing classification;
Hou et al. [19]	218	6	36	(+) functional classification (-) too few classes and less number of models;
Bespalov et al. [9] (NDR - Functional)	70	10	7	(+) functional classification (-) too few models per class;
Jayanti et al. [6] (ESB)	801	42	19	(+) functional classification (+) reasonably sized;
Qin et al. [13]	7464	28	266	(+) functional classification (+) large dataset, suitable for deep learning (-) dataset unavailable;
Koch et al. [12]	1 million	-	-	(+) large dataset & available (-) lack of any category or class information makes it unsuitable for classification;

'+' – Advantages; '-' – Limitations

their functionality. The input CAD mesh models were converted into a histogram representation using enhanced shape distribution, and the extracted representation was then fed into a k-Nearest Neighbor (kNN) classifier. Support Vector Machines (SVMs) were employed for classification approach using surface curvatures as a feature in [18], although the classification was not based on model functionality. The idea of SVMs was also used by [19], where a hybrid of moment invariants, principal moments and geometric ratios were used as input feature vectors. In each of these studies, the classification accuracy was not very high. A comparison of some of the shape signatures has been provided in [6] for classification of CAD models.

B. IMAGES AND 3D GRAPHICAL MODELS
For images and 3D graphical models, there exists a plethora of literature in the last few years that employed advanced deep learning techniques such as Convolutional Neural Networks (CNNs). We mention only a few for the sake of completeness (it may be noted that neural networks have been employed in other fields such as control systems, for example, see [20]). CNN gained popularity starting with AlexNet [21] in the area of image processing, where they employed dropouts that increased the speed of CNN. For further information, please refer to the document on the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) [10]. Further improvements have been made in the network architecture for improving the performance such as VGGNet [22], GoogLeNet [23] etc. [24] gives an overview of the various deep learning

algorithms and architectures available. [25] presents the various challenges that exist in conducting deep learning research while also briefing about the ongoing efforts and future trends of deep learning.

For 3D graphical models, Wu et al. proposed ShapeNet [11], a dataset for volumetric shapes. As in the area of image processing, further improvements were made in the techniques for 3D graphical models either using a point-set representation (as in [26]–[29]) or a Voxel-based representation (as in [30], [31]). While a few machine learning approaches perform well for images (sparse representations, manifold learning etc.), they are not popularly used for 3D data because such methods exploit the structured representation of data. For unstructured 3D data such as point sets etc., they may not perform very well.

C. SUMMARY
Table 1 provides a few details on the existing datasets for CAD models along with the number of models, the number of classes/class along with the category of classification. As can be seen from the table, most of them have only a very few number of models. Only ESB [6], which has 801 models, is available for public use, whereas the one in [13] is not available for public use. From Table 1, it can also be observed that the datasets are either not sufficient enough or not available publicly for a deep-learning based approach. Hence, there is a requirement to generate a dataset consisting of a few thousand models with labeled classification that can then be used for deep learning purpose.

In general, in the field of CAD/Engineering, very few problems have employed deep learning approach. Balu et al. [32] developed a voxel-based 3D CNN approach aimed at a framework for the design for manufacturability (DFM). Recently, Zhang et al. [33] proposed FeatureNet, another voxel-based 3D CNN approach to learn machining or manufacturing features. For the task of classification of CAD models, a first of its kind deep learning approach was described in [13] based on a proprietary dataset (7464 models from 28 categories). Using light field descriptor (LFD) for generating 2D images, they were converted to 1D feature vector using Zernike moments descriptor. This input vector was then fed into the deep neural network (DNN), and classification results were obtained.

It can be clearly seen that the number of works are quite limited in the area of classification of CAD models. Even considering recent times, i.e., the last few years, not much literature is available, to the best of our knowledge. Our aim is to bridge this gap by generating a dataset as well as using even further advances made in deep learning - CNN. Hence, in this paper, CNN, is employed for the functional classification of CAD models, perhaps for the first time. The focus is on the dataset 'CADNET', which combines ESB, NDR as well as newly created 3D models.

III. CLASSIFICATION PIPELINE

The overall pipeline for classifying a 3D CAD model using deep learning can be broadly described as follows:

1) Obtaining a dataset of 3D CAD Models which is suitable for training and testing a deep learning architecture.
2) Extracting features from a CAD model using a feature extraction method.
3) Building a deep learning architecture that can efficiently be trained using the extracted representation as input.
4) Post-processing of results (if any).
5) Testing the network for performance.

An overview of the possible classification pipelines is shown in Fig. 2. In the following sections, each step of the pipeline (Figure 2) is explained in greater detail.

IV. DATASET PREPARATION
A. EXISTING DATASETS
A summary of existing datasets is shown in Table 1. As can be seen, some of them are based on the manufacturing process and not functionality. In the datasets that contain functional classification, most of them are not prepared for deep learning usage and hence contain very few models. The dataset from [13], however, is set up with the exclusive purpose of deep learning usage and hence appears to be very useful. It has 7464 models with 28 categories. The main limitation here is that it is a custom dataset and hence is not available for public usage. Also, the dataset has a high class imbalance, as observed from their paper. More recently, the ABC Dataset [12] containing 1 million CAD objects has been made

available. While it may sound promising, it is a mere collection of CAD objects which lacks any class or category information, thus making it unsuitable for a deep learning-based classification approach.

B. 'CADNET'–A COLLECTION OF 3D CAD MODELS
In the NDR dataset (based on functional classification) [9], there are 70 models over 10 classes. Although insufficient for training a deep neural network in itself, it is publicly available and the classification is based on the functionality. In the ESB dataset [6], there are 801 models over 42 categories (excluding the objects from 'Miscellaneous' class), also classified based on their functionality. ESB is a reasonably sized dataset, contains objects from many categories and is publicly available. The models from these two datasets are collected and have been combined into a single dataset - after checking for duplicates, overlapping classes etc. This resulted in a collection of 868 3D CAD models over 43 categories. Although we obtain a decent average of 20 models per-class, this collection is quite imbalanced with the number of models per category ranging from as low as 4 to the highest being 61. Herein arises a need to generate more data in order to increase the size of the dataset as well as to cover the imbalance as much as possible.

The procedure adopted in order to achieve this is as follows. By observing the 3D objects in each category, an overview of the 3D designs is obtained. Using this knowledge, a few representative models are parametrically designed in Autodesk Fusion360 software - for each class. Following this, more 3D models are generated via a python script linked to the Autodesk Fusion360 API, for various sets of parameter values.

For example, in order to model a cuboid parametrically, one needs three parameters - for the three dimensions (say l, b & h). For different sets of values for l, b & h, we get corresponding cuboids. This process essentially generates multiple variants of a certain category by using many sets of parameter values. In this way, many training examples are created. Every generated 3D model is then verified for correctness. Repeating this process for every category in the collection mentioned above, a dataset of 3317 3D CAD objects over 43 categories is obtained. We refer to this dataset as 'CADNET'. The dataset is made available at https://github.com/bharadwaj-manda/CADNET_Dataset

Table 2 shows the details of CADNET, which gives the category name and number of models in each of them. A few sample models from CADNET are shown in Figure 3. The dataset CADNET now has a significantly large number of models, which can be used in a deep learning setting. The classes are still not balanced, and the method adopted to tackle this is presented in section VII-B.

V. FEATURE EXTRACTION AND CNN-BASED CLASSIFICATION
In the recent past, the emergence of other deep learning techniques has revolutionized the use of machine learning
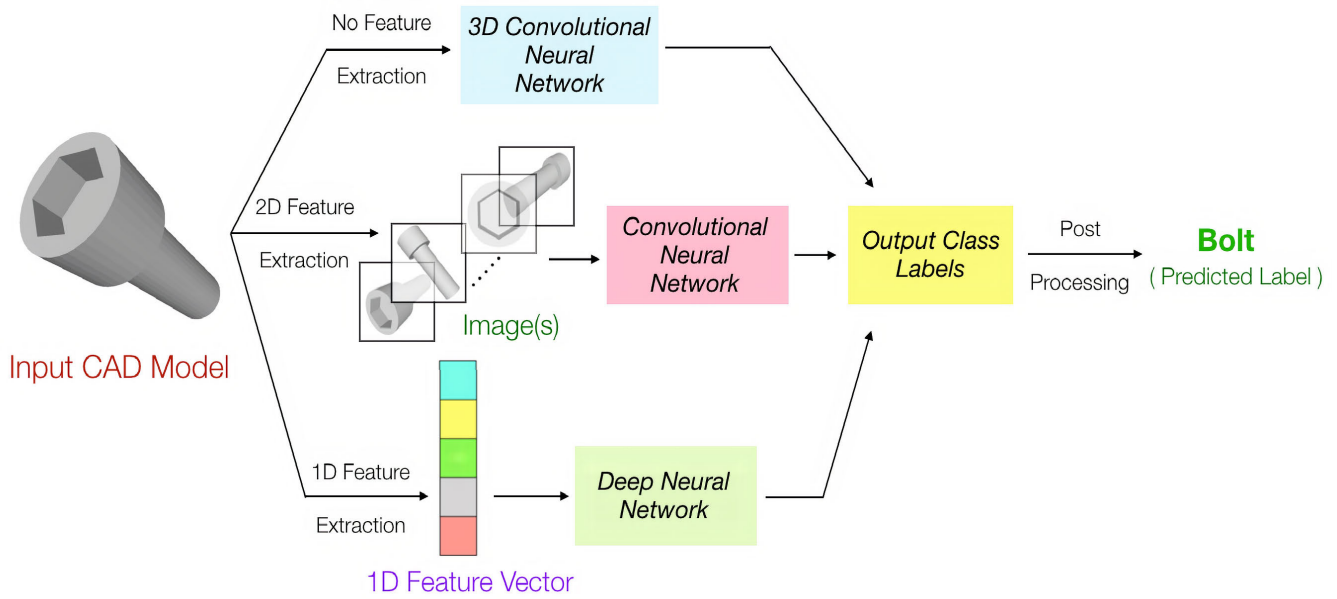
FIGURE 2. Illustrating the possible classification pipelines for 3D CAD Models using deep networks. We use the approach proposed in the 'middle.'

TABLE 2. Details of the developed 'CADNET' dataset - name of the category and the number of models in each of them.

Index	Category Name	# of models	Index	Category Name	# of models	Index	Category Name	# of models
0	90_degree_elbows	100	14	Gear_like_Parts	97	28	Pulley_Like_Parts	61
1	BackDoors	57	15	Handles	119	29	Rectangular_Housings	70
2	Bearing_Blocks	50	16	Intersecting_Pipes	50	30	Rocker_Arms	60
3	Bearing_Like_Parts	50	17	L_Blocks	107	31	Round_Change_At_End	51
4	Bolt_Like_Parts	111	18	Long_Machine_Elements	77	32	Screws	111
5	Bracket_like_Parts	27	19	Long_Pins	104	33	Simple_Pipes	66
6	Clips	54	20	Machined_Blocks	59	34	Slender_Links	60
7	Contact_Switches	60	21	Machined_Plates	99	35	Slender_Thin_Plates	62
8	Container_Like_Parts	60	22	Motor_Bodies	58	36	Small_Machined_Blocks	62
9	Contoured_Surfaces	55	23	Non-90_degree_elbows	108	37	Spoked_Wheels	57
10	Curved_Housings	51	24	Nuts	125	38	Springs	55
11	Cylindrical_Parts	94	25	Oil_Pans	58	39	Thick_Plates	82
12	Discs	163	26	Posts	109	40	Thin_Plates	83
13	Flange_Like_Parts	109	27	Prismatic_Stock	86	41	T-shaped_parts	65
						42	U-shaped_parts	75
							Total # of models	**3317**

for various research domains. A significant breakthrough in the field of image classification has been achieved in [21], where Convolutional Neural Networks (CNNs) are used for the task of classifying more than 1 million images belonging to 1000 classes, as a part of the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) in the year 2012. This motivated us to employ CNN for CADNET.

A. FEATURE EXTRACTION—LIGHT FIELD DESCRIPTOR (LFD) WITH WEIGHTED VIEWS

CNN has been demonstrated extensively on 2D images as inputs. As our inputs are 3D models, it is not very evident on how to convert them to images. Nevertheless, there exists a popular approach called light field descriptor (LFD) that

generates a set of images from a 3D model that can then be used for 2D CNN. The idea of LFD, as described by [34], uses 20 cameras placed at the vertices of a regular dodecahedron to capture images of the 3D model from various views. We use this method to obtain 20 images for every CAD model from our dataset. Each of these images is then assigned the same class label as that of the corresponding CAD model. This process is repeated for every model in the dataset. Now, we use this set of annotated images as the input data for the convolutional neural network.

It should be noted here, however, that this LFD technique is not the same as the one used in [13]. In [13], 10 light fields are created for each 3D model, and 10 images are extracted from each light field. Thus, in effect, 100 images are extracted for
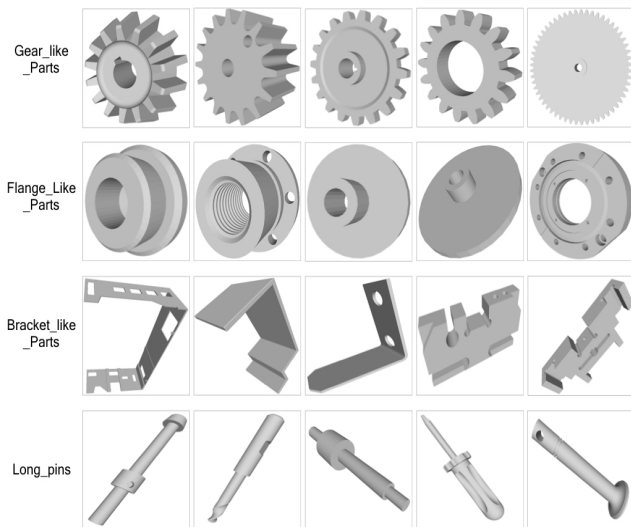
FIGURE 3. Sample models from the 'CADNET' dataset.

FIGURE 4. Illustrating the idea of residual connections [14].

a single 3D Model which are further processed by a Zernike Moments descriptor that produces a 1D feature vector. In our case, we greatly simplify the process by using just 20 images per 3D model (i.e. one light field). The reasons for using only 20 images over 100 images as mentioned in [13] are because (1) Chen et. al [34] showed that 20 images are sufficient and using more than that leads to redundancy. We also observed a similar trend when more number of images were used, and, (2) Using 20 images also takes much lesser time for computation.

On top of that, we employ a post-processing scheme (as will be discussed in Section VI-D), a machine learning algorithm, which learns the influence of each viewing direction on the output prediction. Thus, in effect, we use a weighted LFD approach to extract images from the 3D model, while also knowing the effect of the individual viewing direction.

VI. NETWORK ARCHITECTURE

As part of the ILSVRC challenge, every year newer, deeper, and more efficient architectures have been proposed, outperforming the previous ones. AlexNet [21] was the first architecture to have shown tremendous improvement in classification performance using CNNs when compared to earlier methods. VGGNet [22] has bettered the classification results as compared to AlexNet. Considered to be very deep at the time of its introduction, VGGNet takes an enormous amount of time to train. GoogLeNet [23] (sometimes referred to as InceptionNet v1) provided further improvement in accuracy by using filters with multiple sizes operating on the same level, i.e. using a 'wider' network rather than a 'deeper' one. The winner of the 2015 edition of the challenge is an architecture proposed by [14], popularly known as ResNet. It uses the idea of residual connections, that helps in faster training of deeper networks. Inspired by this idea of residual connections (see Fig. 4), and the advantages it offers for faster and effective training of deep networks, we build a 35-layer
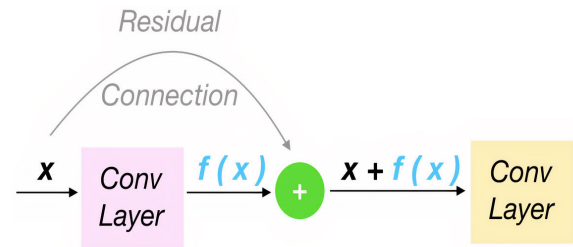
CNN (see Fig. 5). As discussed in Section V-A, 20 images are extracted from each 3D CAD model in the dataset (both for train set and test set). For each of these 20 images, the class label of the corresponding input model is assigned to them. The images from the training set are then used for training the CNN.

A. INPUT LAYER

The input to the network is an image of size 256*256*1. Our input convolutional layer consists of 32 filters, each of which performs a 7*7 convolution operation on the input image. The output of the convolutions is then passed through a series of hidden convolutional layers before reaching the final output layer. The activation function used is ReLU.

B. HIDDEN LAYERS AND RESIDUAL CONNECTIONS

The hidden layers consist of many residual blocks. In general, when we have a series of convolutional layers, the input of each layer is simply the output of the previous layer. However, when we have a residual connection between two layers, the input of each layer will be a summation of the output of the previous layer and the value from the residual connection (See Figure 4 for illustration). The hidden layers are organized as shown in Figure 5. The arrangement of the hidden layers is explained as follows.

There are 5 'groups' of hidden layers (each indicated by a different color in Figure 5). Each group has three residual blocks. Residual connections exist between consecutive residual blocks. Each residual block has two hidden layers, and hence there are six hidden layers per group. Batch normalization is applied at the beginning of every residual block. Solid lines indicate that the number of filters remains the same, while dashed lines indicate that there is an increase in the number of filters by a factor of 2. The number of filters in each layer and the filter size are mentioned across each layer in the Figure. Convolutions are performed using stride = 1. Some hidden layers are indicated by a '/2' in the Figure. In such layers, two additional operations are performed ahead of the batch normalization - (1) max-pooling of size 2*2, (2) 1*1 convolution operation with stride = 2. ReLU activations are used in all hidden layers.

Batch normalization is applied to the output of the last layer of the last group. It is then fed into a pooling layer that performs average pooling of size 4*4. The output is then flattened into a 1D vector and is then fed into a series of two
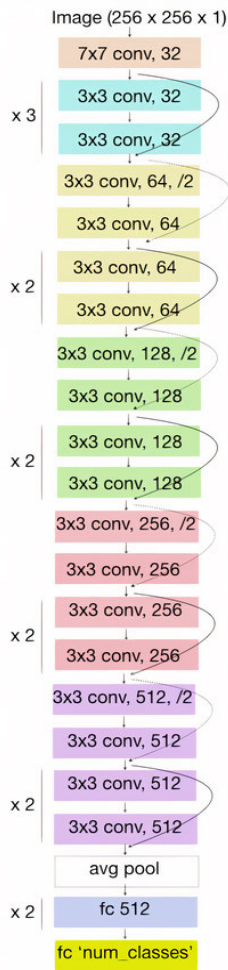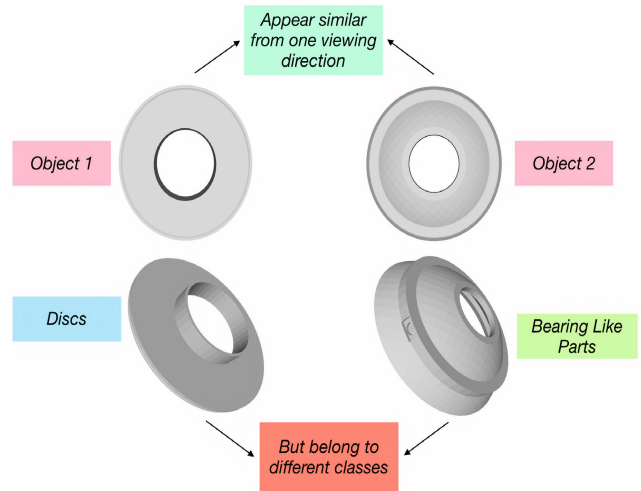
FIGURE 5. Proposed CNN Architecture. This network requires lesser number of filters thereby reducing the number of parameters. Details are presented in Section VI.

FIGURE 6. Need for a post processing scheme to reduce error due to misclassifications.

fully connected layers, each of 512 nodes. ReLU activation is used in both these layers. Dropouts with a probability value of 0.25 are applied in these two layers in order to enhance the prediction accuracy and to avoid any overfitting.

The total number of hidden layers = 5*6 (convolutional layers) + 1 average pooling layer + 2 fully connected layers = 33 layers.

It is to be noted here, that although the idea of residual connections is adopted from [14], the proposed network architecture differs significantly from that of ResNet. These differences arise primarily due to the differences in the nature of the data involved. ResNet, built for the purpose of ImageNet Classification, deals with images that contain real-world graphical objects and each datum consists of many details. In our case, we specifically focus on the images that are extracted from Engineering/CAD Models, which have relatively lesser information as compared to the images from ImageNet dataset. The images are also single-channel, unlike ImageNet, where each image is an RGB (3-channel) image. It should also be noted that although the images extracted from LFD are of size 256*256, it only serves as an outer

boundary for the 3D objects. The actual image is contained within this square, and a significant portion of the image is empty, unlike the images from ImageNet. Hence,

1) The proposed network requires a fewer number of filters to capture the features in the initial layer (32) as opposed to ResNet (64).
2) We require a lesser number of filters (in the hidden layers) compared to ResNet to extract the features from the images. Hence we have six layers of 32, 64, 128, 256 and 512 (= 5952 filters) as opposed to 6 layers of 64, 8 layers of 128, 10 layers of 256 and 6 layers of 512 (= 7040 filters) in ResNet.
3) This, in turn, reduces the number of parameters (by about a million in the network)

C. OUTPUT LAYER

The output from the last hidden layer is then fed into a fully connected layer, with the number of nodes equal to the number of classes. In our case, it is 43 for CADNET. The activation function used here is softmax. The output of this layer is a 1D vector of size 43. The values of this vector indicate probability values. Based on the highest probability value, the class label for the input image is obtained.

D. POST PROCESSING

The output of the network is a vector of 43 probability values for every input image. Generally, the class with the maximum probability value is chosen to be the class label. In this case, however, the machine learning algorithms such as XGBoost [35], and CatBoost [36] are used to do this. The reason is that LFD is a view-based method, and some images are misclassified because models from different classes appear similar from a certain viewing direction (Refer to Figure 6). In order to reduce such misclassifications and thus enhance the prediction accuracy, we process the probability values using XGBoost and CatBoost. These algorithms output a single class label for the image by learning the effect of
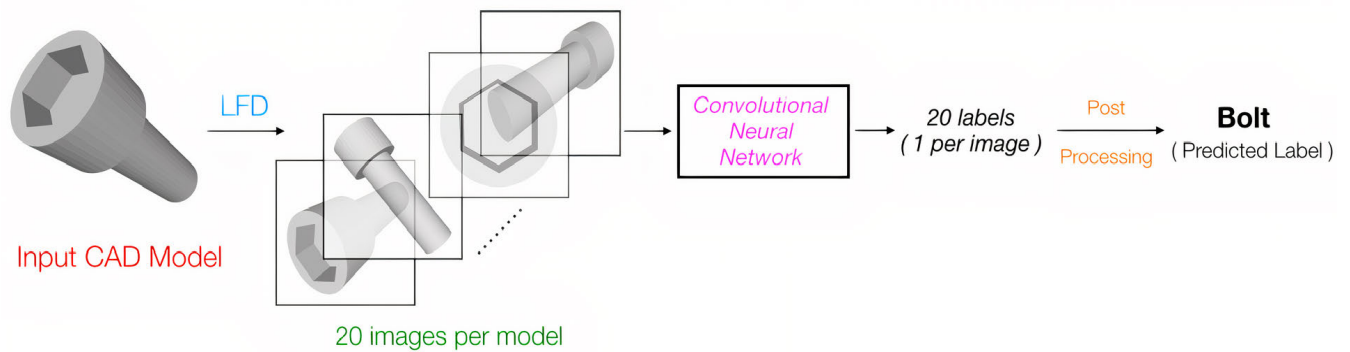
FIGURE 7. Pipeline of our approach for the classification of a given CAD model using LFD, convolution and post-processing (weighting using boosting approaches).

different viewing directions. Now, since each 3D CAD model has 20 images, we would have 20 labels per model. A majority vote of these 20 values is then taken, and a single label per model is obtained.

VII. IMPLEMENTATION DETAILS

The overall pipeline of our implementation for a CAD model classification is depicted in Figure 7 with LFD as the feature extraction, with a CNN using the network architecture as described in Section VI, and with post-processing as discussed in Section VI-D.

A. TRAINING AND TEST SET

In case of large datasets such as [10] for images, [11] for 3D shapes etc., the process of splitting the data into training and test sets is quite straight forward - a percentage of samples are chosen randomly for the test set and the remaining for the training set. This idea generally works because the number of per-class models available for training is high. Also, all the intra-class variations are more or less sure to be captured, owing to the size of the dataset. The same idea is applied for CADNET, with a train-test split of 80-20, which results in 2654 models for training and 663 models for testing. A lower split % for training set resulted in faster training, but at the cost of lesser accuracy - for want of more training data. Higher split % for training set resulted in overfitting the training data, while also taking much longer to train. The 80-20 split used in our training methodology is not arbitrary. It is as per the standard Pareto Principle, which is quite widely used in literature, and our experiments seemed only to reconfirm this. Hence, we presented the results of the 80-20 split, which yields the best classification accuracy.

B. ADDRESSING CLASS IMBALANCE

The problem of class imbalance still exists in this dataset (see Table 2 for details). We tackle this issue by using the idea from [37], during the prediction phase of the neural network. We estimate class weights for the unbalanced dataset. The weight of each class is simply the ratio of the number of

samples to the product of classes and the bin-count of the class labels.

C. HYPERPARAMETER TUNING, LOSS FUNCTION, AND OPTIMIZATION

Training a neural network is a tedious task because of the many decisions involved such as choice of performance metrics, hyper-parameters, loss function, debugging strategies etc. Our choices are mainly based on heuristics ([38]–[40]) and are backed by experimental verification.

After various experiments based on heuristics, we choose a learning rate of 0.001 for training our network. The back-propagation algorithm [41] is used for training the neural network. There are various numerical optimization algorithms available ([42]). We adopt a mini-batch training scheme that uses more than one training example but less than the total number of examples at once. The training examples are split into many batches with 20 examples per batch.

Since the task at hand is multi-class classification, we use the categorical cross-entropy loss function, and the Adam optimization algorithm [43] is used to minimize this loss function. The CNN is trained for 100 epochs. For regularization, there are various methods in practice ([44]) such as enforcing norm penalties, early stopping, etc. In our case, we use the idea of dropouts [45] with a probability value of 0.25 in for the fully connected layers.

D. CODING FRAMEWORK AND SYSTEM CONFIGURATION

For implementing our neural network, we use Python3 with Keras [46] and Tensorflow [47] (gpu-version 1.11.0 [48]). In order to implement the XGBoost and CatBoost algorithms, we use Python3, and sklearn [49].

All the implementations are carried out on a system running Ubuntu 18.04 Operating System. The system has an Intel Core i7-4930K CPU with 64GB RAM and an NVIDIA GeForce GTX 1080Ti GPU with 11GB RAM.

VIII. RESULTS AND DISCUSSION

Our CNN classifier is evaluated for performance on the CADNET dataset, and the results are reported in this section. Training and test sets are chosen, as discussed

TABLE 3. Misclassification results on CADNET.

Index	Category	# of models in test set	# of misclassified models	Misclassified as
0	90_degree_elbows	20	1	Simple_Pipes
4	Bolt_Like_Parts	22	3	Posts, Screws
5	Bracket_like_Parts	5	1	Small_Machined_Blocks
10	Curved_Housings	10	2	90_degree_elbows
11	Cylindrical_Parts	18	1	Posts
12	Discs	32	2	Flange_Like_Parts
20	Machined_Blocks	11	1	Rectangular_Housings
21	Machined_Plates	19	2	Thick_Plates, Thin_Plates
24	Nuts	25	1	Discs
29	Rectangular_Housings	14	2	Thin_Plates, Long_Machine_Elements
32	Screws	22	2	Bolt_Like_Parts
34	Slender_Links	12	3	Long_Machine_Elements, Thin_Plates
36	Small_Machined_Blocks	12	1	Long_Machine_Elements
37	Spoked_Wheels	11	1	Gear_like_Parts
40	Thick_Plates	16	2	Rectangular_Housings, Thin_Plates
41	Thin_Plates	13	4	Rectangular_Housings, Machined_Plates
	Remaining classes	401	0	–
	Total	**663**	**29**	

in Section VII-A. As there exists class imbalance, it is addressed, as discussed in Section VII-B. The obtained accuracy is then put into perspective. Our results are also compared with various other features and their accuracies obtained using a deep neural network.

A. RESULTS ON CADNET DATASET

The CNN training time is approximately 30h, due to a large number of inputs (2654*20 images) that the CNN has to process. The accuracy computed at the output layer of the CNN is 93.41%. The accuracy is further improved using XGBoost/CatBoost algorithm as in Section VI-D. Roughly, the time taken for XGBoost is 15s and for CatBoost is 12s. With XGBoost, the obtained accuracy is 95.63% with 29 models misclassified out of 663 models in the test set. Similar results are obtained using CatBoost - 95.47% with 30 misclassifications. Figures 8a and 8b show the plots of CNN accuracy and test loss respectively with respect to the number of epochs when trained on CADNET. We report the results using XGBoost since it obtains a higher numeric value for accuracy as compared to CatBoost.

Table 3 shows per-class classification results after applying boosting. There are a total of 663 models in the test set, across 43 categories. Out of these, 27 classes are perfectly classified without a single misclassification (401 models); for the majority of the remaining classes, there are just 1 or 2 models that are misclassified. For a majority of these models, the misclassification results are quite understandable since the class predicted by the proposed architecture and the actual class have a lot of similarities. For example, a model from the class Thick_Plates is wrongly identified

as Thin_Plates, since both classes are Plates. Also, 2 models from Bolt_Like_Parts are classified as Screws, and 2 models from Screws are identified as Bolts. A Spoked_Wheel is identified as a Gear and so forth. These classes are pretty similar, and the resulting misclassifications are understandable. This is due to LFD being a view-based technique. For these models, when the extracted images by LFD are visualized, they look very much similar to each other and hence the wrong predictions.
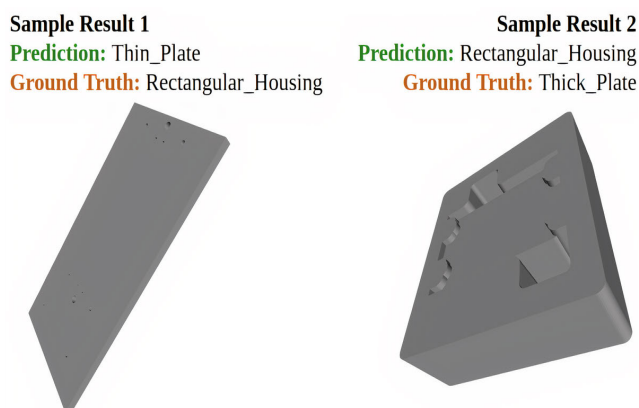
Regarding the apparently non-obvious misclassifications, when the models are visualized, they look very different from the other objects of the same class. In fact, they look similar to some objects of the predicted class. For instance, Figure 9 shows some visual results of the wrongly classified models. The first object is from the class "Rectangular_Housing" which it is misclassified as a "Thin_Plates". Similarly, for the second object, the prediction is "Rectangular_Housing", while the object is a "Thick_Plate". It is easy to see why these misclassifications occur, as the models look very similar to some objects of the predicted class.

B. COMPARISON WITH OTHER METHODS ON CADNET USING DEEP NETWORKS

There is no information available to directly compare the proposed CNN-based approach on CADNET. As there was no deep network-based approach employed on CADNET, we also implemented a deep neural network (DNN). We need to select and extract features from each input 3D CAD model such that the extracted representation can capture the essential information from the input model. As an initial trial, a naive Voxel-based 3D CNN approach is used. It is a 14 layer
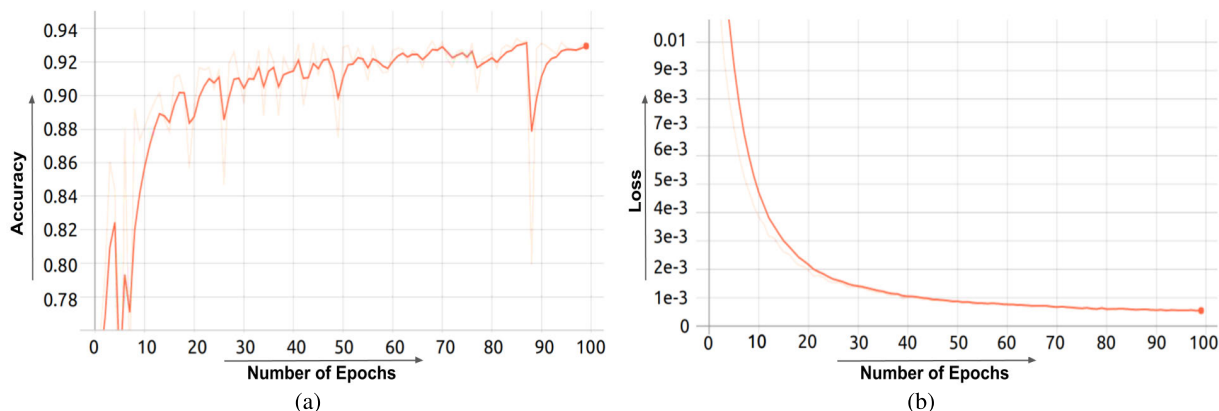
FIGURE 8. Plots of Accuracy and Test Loss over the training period (Number of epochs). (a) CNN accuracy increases as training time progresses; (b) CNN test loss decreases as training time progresses.

Sample Result 1
Prediction: Thin_Plate
Ground Truth: Rectangular_Housing

Sample Result 2
Prediction: Rectangular_Housing
Ground Truth: Thick_Plate

FIGURE 9. Sample misclassification results of the proposed method when trained on CADNET.

network which consists of a series of 3D Convolution and 3D MaxPool layers followed by two dense layers. This naive approach of directly using a Voxel-based 3D CNN, performed very poorly (classification accuracy is 36%), which is to be expected due to the presence of features such as holes etc. in 3D CAD models. Also, the presence of many empty/sparse voxels arising due to sparse nature of the point-sets of CAD models do not help in obtaining a better accuracy.

Further experiments are carried out using an 8-layer, fully connected, deep neural network (DNN). Extensive 3D feature extraction techniques have been proposed in [5], [50], [51] and [52]. For our work, we have tried various geometry-based feature extraction methods such as Euclidean distance between points, geodesic distance, a hybrid of Euclidean and geodesic distances, the angle between normals etc. Using these methods, we obtained a 1D vector representation of the 3D shape and then fed it as an input vector to the deep neural network.

Figure 10 indicates the accuracy results using various feature-classifier combinations on CADNET. It can be observed that the best accuracy obtained is only 84.61% using 3D signatures, while the naive Voxel-based 3D CNN performed very poorly. On the other hand, our method of using LFD with CNN, along with combining weighted

views yielded much higher accuracy. The proposed network architecture is also compared with state-of-the-art CNN architectures such as GoogLeNet [23] and ResNet [14]. The proposed network yielded a maximum accuracy of 95.63%, much higher than the next best one, while also taking much lesser time to train.

C. COMPARISON WITH DEEP LEARNING APPROACHES USED FOR GRAPHICAL MODELS

As mentioned in Section II, most deep learning approaches for graphical models either use a point-set representation or a Voxel-based representation. It may be noted that the point-set representation of CAD models is sparse and non-uniform, thus leading to many empty voxels in the voxel-based representation. Hence, a direct 3D CNN on a voxel-based representation may not work well for the classification task of an entire 3D CAD model (substantiated with results from Figure 10). From this, it can be seen that such methodologies adopted for 3D graphical models need not perform very well on 3D Engineering/CAD models. A few approaches, however, such as the Multi-View CNN (MVCNN) [53], use a view-based method, i.e. use images of the 3D models. Since such an approach appears to work well on CAD models, we experimented with this approach on CADNET.

MVCNN uses two camera setups - (1) 12 views (one view for every 30^0 from 0^0 to 360^0), under the assumption that the input shapes are upright oriented along a consistent axis and (2) 80 views: 20 views that are obtained from viewpoints at the 20 vertices of an icosahedron enclosing the shape and then from each viewpoint, using 0^0, 90^0, 180^0 and 270^0 rotation along the axis passing through the viewpoint and the object centroid. No prior assumption is made regarding the orientation of the object in this case. A TensorFlow implementation of MVCNN with the first camera setup is available publicly on GitHub [54]. The method uses 12 images, extracted from 12 viewing directions, and these images are fed into the network which is trained on CADNET for of 100 epochs (default according to the paper).

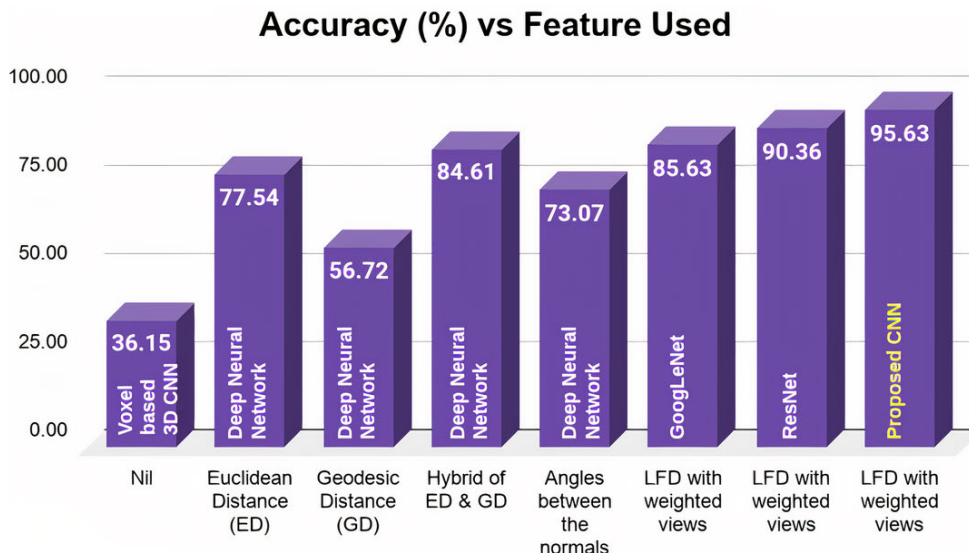The accuracy obtained using this method on CADNET is only 58.75% after training for 100 epochs. By further

FIGURE 10. Comparing various feature-classifier combinations on CADNET.
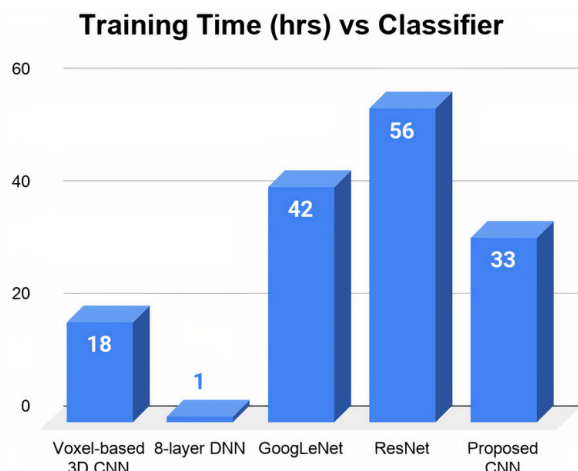

FIGURE 11. Comparing the training time of various networks when trained on CADNET.

increasing the number of epochs, a maximum accuracy of 61.25% is obtained at around 170 epochs, after which no further improvement is noticed. In fact, after 200 epochs, the model begins to overfit the training data, and the test accuracy begins to reduce. This is to be expected due to several reasons. Firstly, MVCNN with the first camera setup uses 12 views, under the assumption that the input shapes are upright oriented along a consistent axis. While this assumption might hold true for datasets such as ModelNet, the same cannot be said for CAD datasets where the objects consist of volumetric features that are not always oriented along a standard axis. Also, using just these 12 views might not be sufficient enough for training. Secondly, the network architecture is not very deep - only 5 conv layers, 1 view pooling layer followed by 3 fully connected (fc) layers (very much like AlexNet which also, incidentally, has 5 conv and 3 fc layers). Considering the fact that deeper architectures for images such ResNet have shown significant improvement in results

as compared to AlexNet, we have also experimented with a modified architecture for MVCNN that is 'ResNet-like'. This resulted in an accuracy of 72.33% - indicating a strong influence of the network architecture on the classification accuracy.

On similar lines, extending the available TensorFlow implementation for the second camera setup, the MVCNN architecture is trained on CADNET once again for 100 epochs. The obtained test accuracy is only 41.25% when trained for 100 epochs. By further training, an accuracy value of 62.5% is obtained at around 260 epochs, beyond which the model begins to overfit the training data. The reason for the failure of MVCNN to perform well on CADNET, even using the second camera setup - without any assumption regarding the orientation of the objects, could be that ModelNet is a relatively well-balanced dataset, with many more number of training examples as compared to CADNET. Since class imbalance is not well-handled in this method, it fails to perform well on CADNET, in which class imbalance is present. Since our method takes into account the imbalanced nature of the dataset, it performs way better than MVCNN on CADNET. In addition to this, there is every possibility that using such a large number of views could lead to overfitting, as not much additional information is obtained through these images. This is evident in [53] where the improvement in accuracy is less than 1%. Like in the first case, we have also experimented with a modified 'ResNet-like' architecture for MVCNN which resulted in an accuracy of 73.67% - once again indicating the strong influence of the network architecture on the classification accuracy.

In summary (see Figure 12), our proposed network architecture, even with plain LFD (without weighted views), performs much better than the highest obtained accuracy using the techniques mentioned above on CADNET, resulting in an accuracy of 93.41% (next highest is 73.67%), despite using a ResNet-like architecture. The reason for this method
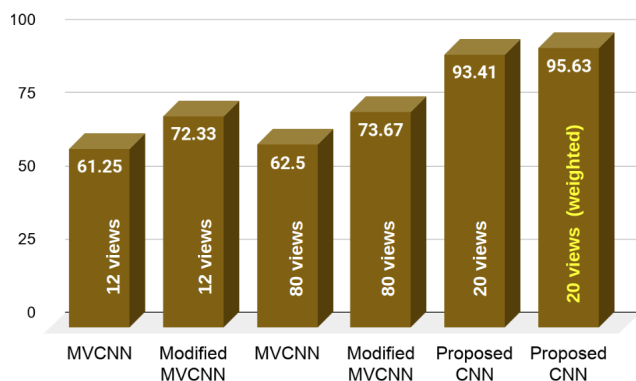
FIGURE 12. Comparing the accuracies of view-based deep learning techniques on CADNET.

to perform well is due to the modifications that have been done to the architecture, to suit CAD model images (as elaborated in Section VI-B). Furthermore, by using the proposed weighted LFD views scheme, the accuracy is improved further to 95.63%. From these experiments, it can be concluded that the proposed network architecture, along with the view-weights and class-weights approach achieves the best performance on CADNET.

D. LIMITATIONS AND POSSIBLE FUTURE WORK

The scope of this work is limited to 3D CAD Mesh models. Other kinds of inputs, such as images or even 3D point sets etc., are not handled by the proposed approach. It is worth exploring to consider building a unified network architecture to process multiple input formats. Also, one possible way to improve the results could be to use a much deeper network and with many more filters in each layer.

Also, when the dataset is made open, users could contribute towards enhancing the dataset, which in turn can be used to increase the performance of the network. As an extension of this work, an automatic CNN-based 3D CAD model retrieval system can be developed. The input queries to the search engine can be compared against the models from CADNET database, using the classification results from the proposed CNN. The current work could also be extended to a CAD assembly model retrieval [4] and sketch-based retrieval of CAD models [55].

IX. CONCLUSION

We built a collection of 3D CAD (Engineering) models with functional classification, termed as CADNET, using the available data from existing datasets ESB and NDR, and augmenting them with manually generated models. A convolutional neural network (CNN) classifier for 3D CAD models was then built, perhaps for the first time. It was observed that processing the 3D CAD models directly into a 3D CNN yielded poor results. Hence, light field descriptor (LFD) was then used for extracting features from a CAD model, and the obtained images were fed into the proposed CNN. A residual network architecture for CAD models with much lesser number of filters (thereby reducing the number of

parameters and the time for training) was also proposed. We also observed that 20 images per CAD model was sufficient. The problem of class imbalance was addressed by using a class-weight approach. Post-processing of the CNN results was done using XGBoost / CatBoost. It was also shown that proposed approach resulted in the highest classification accuracy when compared to other features/networks. Possibilities of extending this work to related research problems have also been discussed.

REFERENCES

[1] T. G. Gunn, "The mechanization of design and manufacturing," *Sci. Amer.*, vol. 247, no. 3, pp. 114–131, 1982.

[2] D. G. Ullman, *The Mechanical Design Process*, vol. 2. New York, NY, USA: McGraw-Hill, 2010. [Online]. Available: https://www.davidullman.com/mechanical-design-process-6ed

[3] J. Bai, S. Gao, W. Tang, Y. Liu, and S. Guo, "Design reuse oriented partial retrieval of CAD models," *Comput.-Aided Des.*, vol. 42, no. 12, pp. 1069–1084, Dec. 2010.

[4] Z. Han, R. Mo, H. Yang, and L. Hao, "CAD assembly model retrieval based on multi-source semantics information and weighted bipartite graph," *Comput. Ind.*, vol. 96, pp. 54–65, Apr. 2018.

[5] N. Iyer, S. Jayanti, K. Lou, Y. Kalyanaraman, and K. Ramani, "Three-dimensional shape searching: State-of-the-art review and future trends," *Comput.-Aided Des.*, vol. 37, no. 5, pp. 509–530, Apr. 2005.

[6] S. Jayanti, Y. Kalyanaraman, N. Iyer, and K. Ramani, "Developing an engineering shape benchmark for CAD models," *Comput.-Aided Des.*, vol. 38, no. 9, pp. 939–953, Sep. 2006.

[7] L. J. Latecki, R. Lakamper, and T. Eckhardt, "Shape descriptors for non-rigid shapes with a single closed contour," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, vol. 1, Jun. 2000, pp. 424–429.

[8] P. Shilane, P. Min, M. Kazhdan, and T. Funkhouser, "The princeton shape benchmark," in *Proc. Shape Model. Appl.*, Jun. 2004, pp. 167–178.

[9] D. Bespalov, C. Y. Ip, W. C. Regli, and J. Shaffer, "Benchmarking CAD search techniques," in *Proc. ACM Symp. Solid Phys. Modeling (SPM)*, New York, NY, USA, 2005, pp. 275–286.

[10] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, "ImageNet large scale visual recognition challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Dec. 2015.

[11] Z. Wu, S. Song, A. Khosla, F. Yu, L. Zhang, X. Tang, and J. Xiao, "3D ShapeNets: A deep representation for volumetric shapes," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 1912–1920.

[12] S. Koch, A. Matveev, Z. Jiang, F. Williams, A. Artemov, E. Burnaev, M. Alexa, D. Zorin, and D. Panozzo, "ABC: A big CAD model dataset for geometric deep learning," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 9601–9611.

[13] F.-W. Qin, L.-Y. Li, S.-M. Gao, X.-L. Yang, and X. Chen, "A deep learning approach to the classification of 3D CAD models," *J. Zhejiang Univ. Sci. C*, vol. 15, no. 2, pp. 91–106, Feb. 2014.

[14] 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.

[15] M. C. Wu and S. R. Jen, "A neural network approach to the classification of 3D prismatic parts," *Int. J. Adv. Manuf. Technol.*, vol. 11, no. 5, pp. 325–335, Sep. 1996.

[16] C. Y. Ip, W. C. Regli, L. Sieger, and A. Shokoufandeh, "Automated learning of model classifications," in *Proc. 8th ACM Symp. Solid Model. Appl. (SM)*, New York, NY, USA, 2003, pp. 322–327.

[17] C. Y. Ip and W. C. Regli, "Content-based classification of CAD models with supervised learning," *Comput.-Aided Des. Appl.*, vol. 2, no. 5, pp. 609–617, Jan. 2005.

[18] C. Y. Ip and W. C. Regli, "Manufacturing classification of CAD models using curvature and SVMs," in *Proc. Int. Conf. Shape Model. Appl. (SMI)*, 2005, pp. 361–365.

[19] S. Hou, K. Lou, and K. Ramani, "SVM-based semantic clustering and retrieval of a 3D model database," *Comput.-Aided Des. Appl.*, vol. 2, nos. 1–4, pp. 155–164, 2004.

[20] Y. Wang, H. Shen, and D. Duan, "On stabilization of quantized sampled-data neural-network-based control systems," *IEEE Trans. Cybern.*, vol. 47, no. 10, pp. 3124–3135, Oct. 2017.

[21] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds. Red Hook, NY, USA: Curran Associates, 2012, pp. 1097–1105.

[22] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *CoRR*, vol. abs/1409.1556, pp. 1–14, Sep. 2014.

[23] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. E. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," *CoRR*, vol. abs/1409.4842, 2014.

[24] A. Shrestha and A. Mahmood, "Review of deep learning algorithms and architectures," *IEEE Access*, vol. 7, pp. 53040–53065, 2019.

[25] X.-W. Chen and X. Lin, "Big data deep learning: Challenges and perspectives," *IEEE Access*, vol. 2, pp. 514–525, 2014.

[26] R. Q. Charles, H. Su, M. Kaichun, and L. J. Guibas, "PointNet: Deep learning on point sets for 3D classification and segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 77–85.

[27] S. Ravanbakhsh, J. G. Schneider, and B. Póczos, "Deep learning with sets and point clouds," *CoRR*, vol. abs/1611.04500, pp. 1–12, Nov. 2017.

[28] A. Garcia-Garcia, F. Gomez-Donoso, J. Garcia-Rodriguez, S. Orts-Escolano, M. Cazorla, and J. Azorin-Lopez, "PointNet: A 3D convolutional neural network for real-time object class recognition," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2016, pp. 1578–1584.

[29] C. R. Qi, L. Yi, H. Su, and L. J. Guibas, "PointNet++: Deep hierarchical feature learning on point sets in a metric space," in *Proc. NIPS*, 2017, pp. 1–14.

[30] D. Maturana and S. Scherer, "VoxNet: A 3D convolutional neural network for real-time object recognition," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2015, pp. 922–928.

[31] C. R. Qi, H. Su, M. NieBner, A. Dai, M. Yan, and L. J. Guibas, "Volumetric and multi-view CNNs for object classification on 3D data," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 5648–5656.

[32] A. Balu, K. G. Lore, G. Young, A. Krishnamurthy, and S. Sarkar, "Adeep 3D convolutional neural network based design for manufacturability framework," *CoRR*, vol. abs/1612.02141, 2016.

[33] Z. Zhang, P. Jaiswal, and R. Rai, "FeatureNet: Machining feature recognition based on 3D convolution neural network," *Comput.-Aided Des.*, vol. 101, pp. 12–22, Aug. 2018.

[34] D.-Y. Chen, X.-P. Tian, Y.-T. Shen, and M. Ouhyoung, "On visual similarity based 3D model retrieval," *Comput. Graph. Forum*, vol. 22, no. 3, pp. 223–232, Sep. 2003.

[35] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. KDD*, 2016, pp. 785–794.

[36] A. V. Dorogush, V. Ershov, and A. Gulin, "CatBoost: Gradient boosting with categorical features support," *CoRR*, vol. abs/1810.11363, pp. 1–7, Oct. 2017.

[37] G. King and L. Zeng, "Logistic regression in rare events data," *Political Anal.*, vol. 9, no. 2, pp. 137–163, Spring 2001.

[38] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning, Chapter—Practical Methodology*. Cambridge, MA, USA: MIT Press, 2016. [Online]. Available: http://www.deeplearningbook.org

[39] Y. Bengio, "Practical recommendations for gradient-based training of deeparchitectures," *CoRR*, vol. abs/1206.5533, 2012.

[40] H. Larochelle, Y. Bengio, J. Louradour, and P. Lamblin, "Exploring strategies for training deep neural networks," *J. Mach. Learn. Res.*, vol. 10, no. 1, pp. 1–40, Jan. 2009.

[41] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 6088, pp. 533–536, Oct. 1986.

[42] S. Ruder, "An overview of gradient descent optimization algorithms," *CoRR*, vol. abs/1609.04747, pp. 1–14, Sep. 2016.

[43] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, vol. abs/1412.6980, 2015, pp. 1–15.

[44] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning, Chapter—Regularization for Deep Learning*. Cambridge, MA, USA: MIT Press, 2016. [Online]. Available: http://www.deeplearningbook.org

[45] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, 2014.

[46] F. Chollet. (2015). *Keras*. [Online]. Available: https://github.com/fchollet/keras

[47] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, and M. Kudlur, "TensorFlow: A system for large-scale machine learning," in *Proc. 12th USENIX Conf. Operating Syst. Design Implement. (OSDI)*. Berkeley, CA, USA: USENIX Association, 2016, pp. 265–283.

[48] Tensorflow-GPU. *Tensorflow Documentation*. [Online]. Available: https://www.tensorflow.org/guide/gpu

[49] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Pas-Sos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2011.

[50] A. Cardone, S. K. Gupta, and M. Karnik, "A survey of shape similarity assessment algorithms for product design and manufacturing applications," *J. Comput. Inf. Sci. Eng.*, vol. 3, no. 2, pp. 109–118, Jun. 2003.

[51] J. W. H. Tangelder and R. C. Veltkamp, "A survey of content based 3D shape retrieval methods," in *Proc. Shape Modeling Appl.*, 2004, pp. 145–156.

[52] R. Osada, T. Funkhouser, B. Chazelle, and D. Dobkin, "Shape distributions," *ACM Trans. Graph.*, vol. 21, no. 4, pp. 807–832, 2002.

[53] H. Su, S. Maji, E. Kalogerakis, and E. Learned-Miller, "Multi-view convolutional neural networks for 3D shape recognition," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 945–953.

[54] T. Lee. (2015). *An Multi-View CNN (MVCNN) Implementation With Tensorflow*. [Online]. Available: https://github.com/WeiTang114/MVCNN-TensorFlow

[55] F.-W. Qin, S.-M. Gao, X.-L. Yang, J. Bai, and Q.-H. Zhao, "A sketch-based semantic retrieval approach for 3D CAD models," *Appl. Math. A, J. Chin. Univ.*, vol. 32, no. 1, pp. 27–52, Mar. 2017.

BHARADWAJ MANDA received the B.Tech. degree in computer science and engineering from the National Institute of Technology Tiruchirappalli, in 2016. He is currently pursuing the Ph.D. degree with the Department of Engineering Design, Indian Institute of Technology Madras (IIT Madras).

Since 2016, he has been a Research Assistant and Teaching Assistant with IIT Madras. His current research interests include deep learning, geometry, computer graphics, and computer-aided design. He is passionate about teaching. He was also a recipient of the AIEEE Topper's Scholarship (from July 2012 to May 2016) and the Half-Time Teaching Assistantship (since July 2016) both from the Ministry of Human Resource Development, Government of India.

PRANJAL BHASKARE was born in Khandwa, Madhya Pradesh, India, in 1996. He received the B.Tech. and M.Tech. degrees from the Indian Institute of Technology Madras. During this time, he worked on various projects like "Classification of 3D Models" and "Malaria Parasite Detection." He also collaborated with Knorr-Bremse to develop different object detection APIs to achieve vehicle autonomy. He is currently a Data Scientist with Reliance Jio, where he is working on the gene sequence analysis and other business problems.

RAMANATHAN MUTHUGANAPATHY is currently a Professor with the Department of Engineering Design, IIT Madras, where he heads the Advanced Geometric Computing Laboratory (AGCL). His research interests include data analysis using machine/deep learning, algorithm development for problems such as reconstruction, Voronoi diagrams, optimal locations for visibility, and virtual/augmented reality. He has received best paper awards in premier international conferences on solid modeling and CAD.

• • •