IMAGE ORIENTATION CORRECTION A REVIEW

*

1st Sandesh Verma
Department of CSE
MANIT
Bhopal,India

2nd Dr. Namita Tiwari
Department of CSE
MANIT
Bhopal,India

3rd Dr.Meenu Chawla
Department of CSE
MANIT
Bhopal,India

4thVishal Singh Bhati
Department of CSE
MANIT
Bhopal,India

***Abstract*—Image orientation task is easy if seen from the human perception but when it comes to machines it becomes a challenging task to correctly predict the correct orientation of the image .Prior works on this topic dealt with handcrafted techniques for image orientation correction but with the advent of new deep learning techniques the task of orientation correction is fast gaining pace with possessing capabilities to not only make the process fast but also to automate it. This paper focuses on bringing up a review of major advancement in the techniques that have taken over time. This paper will be helpful and insightful for researchers, especially for beginners who want to understand the basics and have the overview of the techniques of image orientation correction at a single place.**

***Index Terms*—Review; Image orientation; Deep learning; convolution neural networks.**

I. INTRODUCTION

When it comes to the human eye and it's capabilities of instantly detecting and analysing the orientation of image it takes a fraction of second, which is still a long way to go for the computer vison applications but is greatly supported by the ever emerging deep learning techniques which have been adding optimal speed and accuracy in correcting orientation of images. The following section will talk about the methods adopted in past for the image orientation correction techniques

II. METHODS AND TECHNIQUES

A. Image pre-processing and thresholding technique

One of the previous works on image orientation correction done around 1992 that dealt with orientation improvement in chest images [18] talks about orientation correction in two phases, the first phase involved reformatting images into the standard pixels of size 1K x 1K or 2K x 2K and the second phase involved correcting the orientation of the image to one of the standard positions. Once the unexposed background was removed as a part of first phase the algorithm for orientation correction was applied, the images were either correctly oriented or had 90 degree or 180 degree clockwise and counter clockwise rotations or were either flipped about the y axis giving us a total of eight possible orientations, then a three step analysis was performed in which the first step involved locating the spine region in image, in the second step on a thresholded image upper extremities or neck region is searched for then orientation along perpendicular direction of the spine is found out and after few more analysis in the final steps decision is made if the orientation is correct or it needs a flip around y axis or not then after combining all these results the orientation angle is found.

In short to sum up the author in this talked about an algorithm that would predict the orientation of the image from one of the standard positions and after following the two phase process one of the eight angles is chosen to rotate the image. This process helped the radiologist to save a lot of time in manually rotating the images.

B. Based on low level visual content

In the research carried out in paper[22] the author focuses on laying importance to low level features [9, 17] rather than on semantics of object and since to efficiently classify the image orientation collection of various visual features would perform better so the author took structural features to capture shape and texture related features and chrominance feature to extract color related features.

The use of combination of both the features help achieve 5% better results as compared to use of any one of the features. In order to determine regional features the author divided the image into N X N sub blocks and then features were extracted from these sub blocks, the value of N was decided empirically by the author to overcome curse of dimensionality [7] .

The chrominance feature were taken only for the peripheral blocks as orientation of the image was better predicted from the periphery rather than from centre portion and for the luminance feature help of EDH (edge direction histogram) were taken and the use of canny edge detector [2] is employed for detecting edges in the image. Both the above stated features were normalized and were fed to the rigorous classification system framework proposed by the author which used mostly the support vector machine classifiers (SVM) [21] The above framework then corrected the orientation of image with better accuracies then the previous known models.

C. Based on Borda Count

The orientation correction technique discussed in [15] talks about the use of two new set of features in addition to the already existing ones which as discussed in [22], the two new low level features that were introduced by the author were (HCH) Harris corner histogram and (PHS) phase symmetry were in HCH the corner features were taken help to capture information present in blocks by counting the number of corner points in every block and for their detection help of Harris corner detector [10] was taken.

PHS was a method employed to extract illumination and contrast invariant features from the image and then the number of symmetric pixels were counted in each block since the level of symmetricity or asymmetricity could be extracted from the invariant features. Now after extraction of features is done it was fed to a multiclassifier system in which various rules could be adopted for final decision from the ensemble of all those classifiers most common among them being voting rule, mean rule, maximum rule, Borda Count [11], minimum rule, of which the best classification accuracy was achieved through the use of Borda Count which is basically a voting technique in which a collection of individual ranking to a combined ranking giving rise to the most relevant decision, the class who is ranked first gets the most points ,the class who is ranked second gets fewer points and so on .

The class with the most points overall wins the election. The author further talks about adopting a rejection rule [22] for images which have a confidence score below a certain threshold .The following table shows the best accuracy obtained by Borda Count when adopted as a combination rule while classification.

TABLE I
ACCURACY OBTAINED CORRESPONDING TO DIFFERENT COMBINATION RULES [9].

Combining rule	Accuracy
Min rule	0.542
Max rule	0.512
Mean rule	0.608
Vote rule	0.598
Borda Count	0.62
DCS	0.541

D. Image Pre Filtering Techniques

The study conducted in [20] on users stated that they feel more stressed on false positive then on false negatives in automatic image orientation correction. Thus the author tries to remove images which could be a potential contributor to the false positives and thus reducing the stress levels of the users. The conventional methods used a 3 class SVM classifier after the extraction of low level features[16] by using a confidence measure for each orientation and summing three against one classifier scores ,the one with the highest confidence is selected if the measure is high enough else rejected

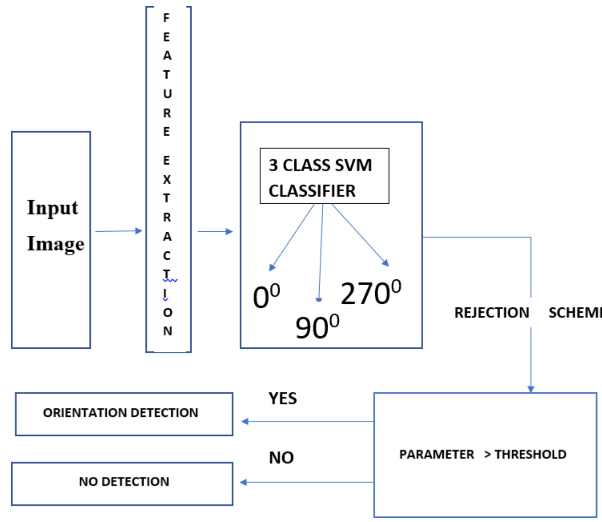

Fig. 1. Common architecture employed for low level feature.

Now the rejection scheme often used to fail according to the author when images were similar to the training data with unsymmetrical composition which made the confidence score high enough to be accepted and surpass the threshold values such images were kept in the category of irrelevant by the author. To overcome this dilemma author proposed a pre filtering technique to filter out irrelevant images and to improve user stress levels.

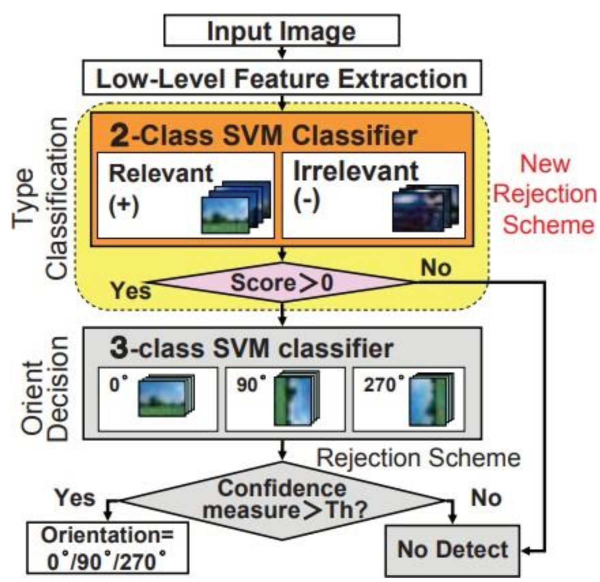

Fig. 2. Use of pre filtering technique for orientation correction [20]

The introduction of a 2 class SVM classifier into the conventional architecture was trained by a bunch of handpicked relevant as well as irrelevant images then by the positive and negative judgement of the classifier it was judged weather the image was relevant for orientation detection or not only the relevant images were fed to the 3 class classifier and not relevant images were unchanged in the process. With the use of this method author was able to reduce the stress levels of the users to a satisfactory levels.

E. Using PCA

PCA stands for principal component analysis which is also one of the feature extraction techniques that can be utilised for image orientation task, the role of PCA is to reduce the dimensionality of the dataset [19] along with generating eigenvectors that points to the direction were the dataset varies the most, every eigenvector produced has a unique value associated with it which is the eigenvalue with the help of which direction of rotation of the new axis can be determined. Now for the task of image orientation using PCA the process followed by the author [4] talks about pre-processing of the image that will involve resizing of image to certain pixel format so that uniformity is maintained then after followed by converting the image to grayscale values after which image is smoothen and noise is removed using Gaussian method and kernel filters followed by binary thresholding of the grayscale image [5].With the help of the canny operator [6] edges are detected then in the final step contours are extracted from these edges for which individual eigenvectors are calculated and further the task of orientation correction is achieved

F. Using Convolutional Neural Networks

A neural network is a collection off interconnected stacks of neuron, just like the biological neuron the artificial neuron is capable of performing calculations. This network is capable of learning patterns by adjusting the weights as the training process is carried forward.

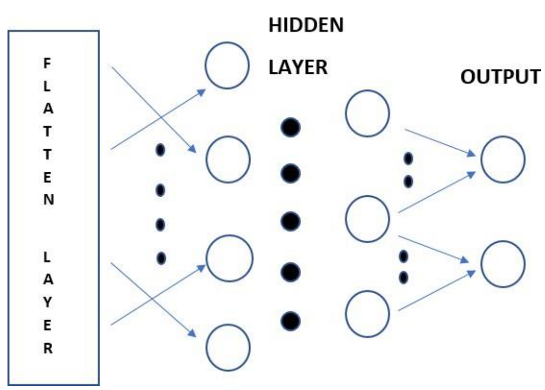

Fig. 3. Basic architecture of neural networks

CNN acts as a special kind of offset of the neural networks that holds the power to process image in an efficient manner, they use convolution operations to extract the features of the image which intern make use of kernel to achieve the task. Kernels are N X N matrix that are run over the image in a sequence to generate a feature map .

CNN also imbibes a pooling operation which helps it to reduce the dimensions of the input feature map and also helps in reducing the amount of computation time and parameters to be tracked. A simplified architecture of CNN is shown in figure 4 for overview of the process that a CNN follows.

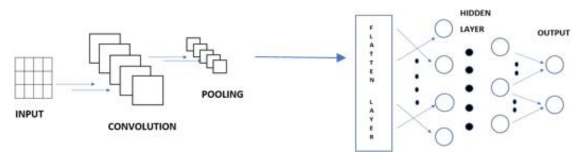

Fig. 4. Simplified view of CNN

Many of the orientation task in past have dealt with document analysis [1, 3, 14] which took into account text layout in lines and shape of letters, in general cases this task becomes cumbersome when image boundaries and other dominant features such as horizontal or vertical lines are missing, to address this issue author [8] proposes a method involving the use of convolutional neural networks, for example in the figure 5 other conventional methods would fail to correctly orient the image where the model given by the author performed better.

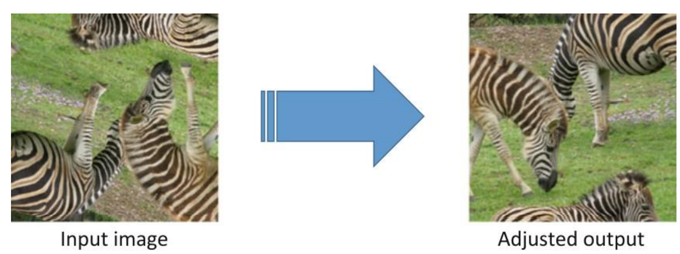

Fig. 5. orientation correction using CNN for the case when horizontal and vertical lines are missing or are misleading [8]

To address this issue author divided the task into three difficulty levels, first one being easy and dealing with images that had at most +30o rotation and the other task being tough and dealing with images having rotations +45o and +360o and called the networks trained on these task as Net-30,Net-45 and Net-360 respectively. For all the task AlexNet architecture [13] was implemented in Caffe and pretrained on ImageNet. This model had 5 convolutional layers 3 fully connected layers and relu activation function after each fully connected layer.

The author took Microsoft COCO training dataset and applied rotations to it while eliminating images with undefined orientation from the training set, also manual verification of the test data set was done to check if the images in test dataset had correct orientation. The author then compares the proposed method with the other state of the art methods and finds that the proposed method outperforms the other methods the results of which can be seen in figure 6.

Task	Net-30	Net-45	Net-360	Net-rough+45	Hough-var	Hough-pow	Fourier
±30°-all	**3.00**	4.00	19.74	19.64	11.41	10.62	10.66
±30°-easy	**2.17**	2.83	19.48	17.90	8.44	7.04	8.64
±30°-hard	**4.26**	5.75	20.12	22.24	15.88	15.99	13.69
±45°-all	-	**4.63**	20.64	19.24	16.92	13.06	16.51
±45°-easy	-	**3.24**	21.26	19.29	14.08	9.01	13.32
±45°-hard	-	**6.71**	19.70	19.15	21.16	19.13	21.28
±180°-all	-	-	**20.97**	**18.68**	-	-	-
±180°-easy	-	-	**20.29**	**18.70**	-	-	-
±180°-hard	-	-	**21.98**	**18.65**	-	-	-

Fig. 6. Average absolute errors of the estimated angles [8]

Some other recent technical works (HP,Technical Disclosure Commons,2022) also talks about use of convolution neural networks to achieve image orientation correction ,the author proposes a method to automatically detect the rotated angle and then reverse the image in opposite direction with the same rotation angle value to correct its orientation mostly for document images. The task was achieved by training the classifier that classifies input images to their respective degree of rotation. Training samples were the images that were given random rotations. Then using CNN predictions were made for the orientation of an image.

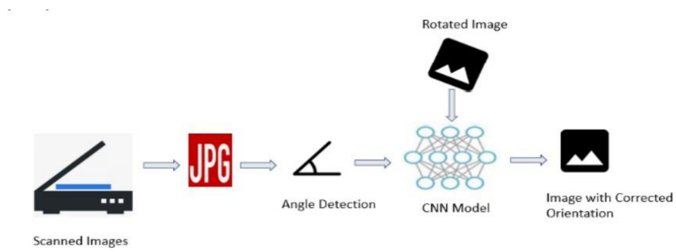

Fig. 7. orientation correction using CNN (HP,Technical Disclosure Commons,2022)

Conclusion

On studying various developments in the domain of image orientation correction we analysed that most of the early works done on image orientation correction mainly focused on extracting low level features, then as the researched progressed more many works started using ensemble methods or multiclassifiers along with the low level feature extraction which helped the model in achieving better accuracy in terms of orientation correction as the study further progressed and with the advancement in the field of deep learning several techniques evolved of which CNN turned out to be the best for the orientation task as it offered many advantages such as making the process automatic also allows us the freedom to generate huge number of training samples so that the network is able to learn subtle contextual features that allow it to estimate the correct orientation. It also provides us the facility to run the network on real time video stream and correct the orientation instantly.

It has been found that the machine learning models are better suited for solving the object-orientation problem in comparison to Principal Component Analysis, hence most of the latest work on orientation correction is being carried out via the use of deep learning framework that is CNN. CNN internally used object detection technique for the image orientation detection [12] which was found out by using guided backpropagation technique in one of the works [12]. It was also found out that using transfer learning was useful for image orientation detection. So it can be concluded that the task of orientation detection possess several scope of improvement while imbibing deep learning techniques and further research on the domain would help it to achieve better accuracies and faster processing time for the orientation task.

References

[1] Bruno Tenório Ávila and Rafael Dueire Lins. "A fast orientation and skew detection algorithm for monochromatic document images". In: *Proceedings of the 2005 ACM symposium on Document engineering*. 2005, pp. 118–126.

[2] JF Canny. "A computation approach to edge detection". In: *IEEE Trans. Pattern Anal. Mach. Intell.* 8.6 (1986), pp. 670–700.

[3] Su Chen and Robert M Haralick. "An automatic algorithm for text skew estimation in document images using recursive morphological transforms". In: *Proceedings of 1st International Conference on Image Processing*. Vol. 1. IEEE. 1994, pp. 139–143.

[4] Amila De Silva. "Object Orientation Detection and Correction Using Computer Vision". In: (2020).

[5] Amila De Silva. "Object Orientation Detection and Correction Using Computer Vision". In: (2020).

[6] Lijun Ding and Ardeshir Goshtasby. "On the Canny edge detector". In: *Pattern recognition* 34.3 (2001), pp. 721–725.

[7] Richard O Duda, Peter E Hart, et al. *Pattern classification and scene analysis*. Vol. 3. Wiley New York, 1973.

[8] Philipp Fischer, Alexey Dosovitskiy, and Thomas Brox. "Image orientation estimation with convolutional networks". In: *German conference on pattern recognition*. Springer. 2015, pp. 368–378.

[9] Borko Furht. *Handbook of multimedia computing*. Vol. 5. CRC Press, 1998.

[10] Chris Harris, Mike Stephens, et al. "A combined corner and edge detector". In: *Alvey vision conference*. Vol. 15. 50. Manchester, UK. 1988, pp. 10–5244.

[11] Tin Kam Ho, Jonathan J. Hull, and Sargur N. Srihari. "Decision combination in multiple classifier systems". In: *IEEE transactions on pattern analysis and machine intelligence* 16.1 (1994), pp. 66–75.

[12] Ujash Joshi and Michael Guerzhoy. "Automatic photo orientation detection with convolutional neural networks". In: *2017 14th Conference on Computer and Robot Vision (CRV)*. IEEE. 2017, pp. 103–108.

[13] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. “Imagenet classification with deep convolutional neural networks”. In: *Communications of the ACM* 60.6 (2017), pp. 84–90.

[14] HK Kwag et al. “Efficient skew estimation and correction algorithm for document images”. In: *Image and vision computing* 20.1 (2002), pp. 25–35.

[15] Alessandra Lumini and Loris Nanni. “Detector of image orientation based on Borda Count”. In: *Pattern Recognition Letters* 27.3 (2006), pp. 180–186.

[16] Jiebo Luo and Matthew Boutell. “Automatic image orientation detection via confidence-based integration of low-level and semantic cues”. In: *IEEE transactions on pattern analysis and machine intelligence* 27.5 (2005), pp. 715–726.

[17] Alexander P Pentland, Rosalind W Picard, and Stan Scarloff. “Photobook: Tools for content-based manipulation of image databases”. In: *Storage and retrieval for image and video databases II*. Vol. 2185. SPIE. 1994, pp. 34–47.

[18] Ewa Pietka and HK Huang. “Orientation correction for chest images”. In: *Journal of Digital Imaging* 5.3 (1992), pp. 185–189.

[19] Fengxi Song, Zhongwei Guo, and Dayong Mei. “Feature selection using principal component analysis”. In: *2010 international conference on system science, engineering design and manufacturing informatization*. Vol. 1. IEEE. 2010, pp. 27–30.

[20] Yusuke Takahashi and Ryoma Oami. “Relevant image pre-filtering in automatic image orientation correction”. In: *2010 IEEE International Conference on Image Processing*. IEEE. 2010, pp. 2657–2660.

[21] Vladimir N Vapnik. “An overview of statistical learning theory”. In: *IEEE transactions on neural networks* 10.5 (1999), pp. 988–999.

[22] Yongmei Michelle Wang and Hongjiang Zhang. “Detecting image orientation based on low-level visual content”. In: *Computer Vision and Image Understanding* 93.3 (2004), pp. 328–346.