OPEN ACCESS

Kashf Journal of Multidisciplinary Research

Vol:01 Issue11 2024

P-ISSN: 3007-1992 E-ISSN: 3007-200X

https://kjmr.com.pk

SUBCELLULAR STRUCTURES CLASSIFICATION IN FLUORESCENCE MICROSCOPIC IMAGES

Saima Batool
Institute of Computing, Muhammad Nawaz Sharif University of Agriculture, Multan, Pakistan.

Ghazanfar Ali
Institute of Computing, Muhammad Nawaz Sharif University of Agriculture, Multan, Pakistan.

Inzamam Shahzad
School of Computer Science and School of Cyberspace Science, Xiangtan University, Xiangtan, Hunan, China.

Muhammad Hanif Soomro
Department of Information Technology, University of Mirpur Khas, Sindh, Pakistan.

Salahuddin
Department of Computer Science, NFC Institute of Engineering and technology, Multan, Pakistan.

Muhammad Ijaz
Department of Computer Science, The Islamia University of Bahawalpur, Pakistan

Article Info

Received: 04th Nov, 2024
Review 1: 06th Nov, 2024
Review 2: 08th Nov, 2024
Published: 10th Nov, 2024

Abstract

Evaluation of cellular protein localization is becoming progressively important. Even after the whole human genome has been sequenced, numerous years will be required to study the structure, function, and each protein localization. Localization information is very important because it provides a context for a protein's structural and functional information. For inssstance, two proteins that possess comparatively similar function and structure may in fact be found in different compartments within the cell hence, these might be included in unrelated cellular processes. Localization data for several proteins were collected using fluorescence microscopy. The resulting patterns were described using a variety of numeric features including local binary pattern (LBP), Haralick's texture features, edge histogram descriptor (EHD), Segmentation-based Fractal Texture Analysis (SFTA) and after that these features were combined to form a single feature vector. This feature vector is then given as input to the classifier to classify the image into 10 different classes for 2D Hela dataset and into 5 classes for CHO dataset. Support vector machine (SVM) with chi-square kernel is used to solve the classification problem. Proposed frame-work obtained 98.9% overall accuracy on Hela dataset and 99.6% on CHO dataset. The obtained accuracy is 0.9% higher on 2D Hela dataset and 0.6% higher on CHO dataset as compared to highest computed accuracy.

Keywords: *Fractal Texture Analysis, Support Vector Machine, local binary pattern*

Introduction

A vital piece of the portrayal of a protein is the assurance of the subcellular organelles or structures to which it limits. This data is profitable because it gives a setting to the protein's structure and capacity. For instance, two proteins that are conjectured to have comparative structure and capacity may in truth restrict to various com-apartments inside the cell and in this manner, be included in unmistakable cell forms. The most well-known strategy for deciding subcellular area is understanding of fluorescence magnifying instrument pictures, both of cells recolored with monoclonal antibodies against an endogenous protein or of cells communicating a GFP-labeled protein from a transfected develop. As of now the understanding is performed outwardly by the specialist. Such subjective understandings might be impacted by examiner predisposition can't be effortlessly affirmed by different agents, don't loan them-selves to factual investigation, and don't give a methodical depiction that can be entered in database.

A mechanized framework for deciphering pictures of confinement example would in this manner have number of points of interest over current practice. These would incorporate objectivity, dependability, and repeatability. We have attempted to create and test techniques for quantitatively depicting such examples [3]. According to the research that has been done, almost everyone would agree that there is currently no collection of work that has been dedicated to providing a quantitative description of where proteins are found. There is no guarantee that biologists have no interest in characterizing the locations of proteins, but this is not often the case. On the other hand, it is frequently necessary to define how a newly discovered protein will be found inside of a cell. Regrettably, at this time, these descriptions are subjective, and as a result, they are not necessarily comparable between different investigators. When assigning a localization pattern to a protein, there is not a standard set of classifications that can be used. This is true even if the investigator's bias could be reduced to a minimum.

In point of fact, due to the complexity and variety of protein localization pat-terns, it is impossible to construct a set of categories that are straightforward enough to let investigators to make subjective assignments of patterns using those categories. As a result of this, the localization of a new protein is typically explained in broad terms, and it is typically done so within a specific sub-domain of cell biology. By consulting the Swiss-PROT database, one can learn about these issues as well as a variety of others. Annotated protein sequences can be found in the Swiss-PROT database. This database provides information on protein structure, function, post-translational modification, variations, and more. In addition to that, one of the fields is devoted to the subcellular localization. When the localization field of the Swiss-PROT database is analyzed by determining the frequency of each distinct localization term1, a number of issues become apparent. To begin, there is a lack of a systematic approach to the descriptions of the localization. The person who entered the information into the database has added their own subjective commentary to many of the entries.

The principal goal of this research work is to display practical use of CBIR system for medical experts. The focus of the research is on the use of various feature extraction techniques in classification. Followings are the most important objectives of the proposed research work.

- Experiment with various feature extraction techniques.
- Experiment with hybrid feature vector, which is obtained by combining the output of different feature extraction techniques.
- Propose a model for classification for medical images, which offers high ac-curacy as compared to the existing ones.

2 Literature Review

Fluorescence microscopy, pattern recognition, and machine learning are three of the fields that have contributed to the development of the technology that has recently been used to build applications such as sub-cellular protein localization. Other fields that have contributed to

the development of this technology include electron microscopy, computational learning, and pattern recognition. This research aims to establish methodologies that will enable the numerical description and subsequent classification of the patterns that can be seen in fluorescent light microscope images of cells. The research was funded by the National Institutes of Health (NIH) and is being funded by the National Science Foundation (NSF). After labeling one or more sub-cellular structures with fluorescent dyes, these images are acquired by capturing images of the ensuing pattern of fluorescence using a microscope. The labeling of the sub-cellular structures can be done in either direction. This results in the difficulty of characterizing these patterns in a manner that is amenable to additional processing, which in turn results in the difficulty of actually acquiring the images themselves [10].

The process of automated sub-cellular localization has quite a few positive applications and benefits. The most essential benefit is that a previously impossible level of standardization may now be achieved through the use of quantitative descriptions of images. A rapid comparison of a new pattern with a large number of already produced patterns in the database could be of tremendous potential use. The same can be said for the study of the interactions between proteins. If we had access to a system like this one, we would be in a far better position to get an understanding of the intricate mechanisms that govern protein interaction and localization.

Identification of proteins with subcellular localization by computational means corresponds to a classification challenge. Because the search space for classifiers that model each individual pixel is too large for practical purposes, machine learning approaches typically do not classify the phenotypes directly from the raw image data. This is because all of these approaches are affected by the Curse of Dimensionality [14], which states that these approaches all share a common flaw. Because of the complexity of these algorithms, we must construct a much more limited number of functions that operate on the raw input image. The outputs of these functions are the only ones that are taken into consideration when classifying the image. Extraction of features is the term used for this process. Applying dimensionality reduction methods and selecting the dimensions that contain the most relevant information is another option. This is erred to as feature selection. Typically, feature selection is used in tandem with feature extraction [43]. Feature extraction is referred as the input pixel co-variance has structure [44], and certain functions of the input are a priori known to not hold discriminative power. Sometimes customized characteristics are utilized, and these features may be based on a numerical method (like skeletonizing), or they could be based on additional information provided by the experimental setup. Both of these options are possible (intensity overlap with an additional marker) [45]. Murphy et al. outlined and continue to update a collection of subcellular localization features (SLFs) that is comprised of multiple categories of picture features. They have been making almost all of their classification decisions based on this foundational set for describing protein spatial distributions [46].

The discipline of machine learning includes a vast array of algorithms that may predict class labels by making use of a labelled set of instances (training set) that is supplied by an experienced individual (Supervised Learning Approach). Among them, the following are some of the ones that have been taken into consideration for the classification of different cell phenotypes: Neural networks [46], Support vector machine (SVM) [48, 49]. The task of classification is difficult due to a number of different facets of the problem. First, there is a huge number of alternative localizations (more than twenty), yet classifiers are normally created using a binary classification framework. For instance, the SVM method involves finding the hyper-plane that provides the best separation between two different sets of vectors in an enhanced space that was created by utilizing kernel functions [50].

3 Methodology

The application of computer vision techniques to the image retrieval problem, also known as the difficulty of searching for digital images in huge databases, is what is known as content-based image retrieval or CBIR for short. The fundamental applications of CBIR are medical analysis, research, and teaching. A substantial advantage originates from the territory of educating. Here instructors can utilize tremendous storages to locate some fascinating cases to show to students. These selected cases are Chosen cases are shown and construct in light as well as outwardly comparative images are introduced to improve the instructive abilities. And huge archives are also utilized by medical students for educational purpose. Another advantage is originated from research territory; the researcher can incorporate their case in their explorations. At last, the most critical and troublesome application region of CBIR is diagnostic; this leads to the combination of the system to the daily life. Content-based image retrieval (CBIR) system allows perusing the database and finding analysis based on visual resemblances. In this chapter, we are going to discuss proposed methodology and different techniques used in our framework. The proposed framework consists of several stages which include feature extraction, feature vector construction, dimensionality reduction and training of classifier and finely the modality classification for the given query image. Fig 1.

3.1 Feature Extraction

The data can be changed into a more limited set of characteristics if an algorithm's input data is too extensive to be processed and there is a good likelihood that it contains redundant information. In this case, the data will be transformed (features vector). This method is often erred to as the "extraction of features" technique. It is hoped that the chosen characteristics will include relevant information derived from the supplied data. Because of this, it will be possible to carry out the activity that was planned by making use of this condensed representation rather of the complete original data.

Features for a digital image are distinct quantifiable properties. The selection of separating and autonomous features is a vital step for a successful CBIR system, the type of the features is generally numeric; however basic components, for example, strings and diagrams are utilized as a part of recognition and retrieval frameworks. The idea of feature correlated to the controlled variable in the linear regression. The process of mining and selection of a feature is a blend of skill and science, the development of such framework to do so is known as feature engineering [24]; it requisite several experiments with various possibilities and grouping of many automated techniques and domain knowledge. Mechanizing this process is known as feature learning [25] where machine learns the feature by itself. In our framework, we have used a set of features, our methodology can be portrayed as scaling up to use as many features and data as possible. The experiments exhibit, that expanding either axis tend to increase performance. Features used in our framework are Haralick’s Features [4], local binary pattern (LBP) [26], edge histogram descriptor (EHD) [27], Segmentation-based Fractal Texture Analysis (SFTA) [28].

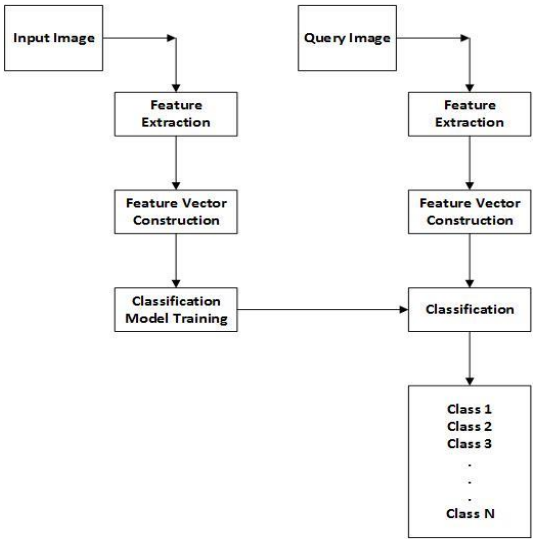

Fig 1: Block Diagram of the Proposed Technique

3.2 Local Binary Pattern (LBP)

The fundamental thought behind the Local binary pattern (LBP) [23] methodology is to utilize the data about the surface from a nearby neighborhood. To begin with, we specify a radius R. Following this, the algorithm creates a binary pattern or code that represents the nearby texture in the area set of P pixels. The binary pattern is generated by taking into consideration the center value as a threshold value. The obtain value is then converted to decimal by using Equation 1-2.

$$LBP(x_c, y_c) = \sum_{n=0}^{7} m(g_n - g_c)2^n \quad (1)$$

$$m(k) = \begin{cases} 1, & k \geq 0 \\ 0, & k < 0 \end{cases} \quad (2)$$

Where g_c represents the value of the center pixel and g_n represents the value of eight surrounding pixels and function m (k) returns a binary value. The length of LBP histogram depends on several neighbors. We utilized uniform pat-tern for experiments because it reduces the size of the histogram by combining the entire non-uniform pattern into a single bin. We experimented with the radius size of 1, 2, and 3 and setting the neighborhood size to 8, 16 and 24 respectively.

3.3 Edge Histogram Descriptor (EHD)

The edge histogram descriptor (EHD) quantifies the degree to which an image's edges are dispersed locally. When the primary texture is not consistent, edge dispersion is a useful mark for comparing images since it shows the variation in the texture. EHD first isolates the image into 4x4 grids, and then specifies the edge in each and every one of those sub images. This is how it determines the edge. The edges that were extracted from the sub-image are sorted into the following five categories: vertical, horizontal, 45-degree, 135-degree, and non-directional. The existence of an edge in each category results in the development of a histogram bin, which results in a histogram with a total of 80 bins.

An interesting variation of EHD is to compute an extended histogram by making use of a histogram that has already been retrieved and partitioned into 80 bins. By mix-ing the image blocks, it is possible to extend the capabilities of an 80-bit histogram. A global histogram, which is created by putting all 16 picture blocks together, and a semi-global histogram are the names given to the extended bins (formed by pooling by image blocks four rows and four columns and five groups). This result is presented in five bins for the global histogram as well as for the semi-global histograms that are derived from the eighty local histogram bins. Therefore, the total number of bins comes to 150.

An interesting variation of EHD is to compute an extended histogram by making use of a histogram that has already been retrieved and partitioned into 80 bins. By mixing the image blocks, it is possible to extend the capabilities of

an 80-bit histogram. A global histogram, which is created by putting all 16 picture blocks together, and a semi-global histogram are the names given to the extended bins (formed by pooling by image blocks four rows and four columns and five groups). This result is presented in five bins for the global histogram as well as for the semi-global histograms that are derived from the eighty local histogram bins. Therefore, the total number of bins comes to 150.

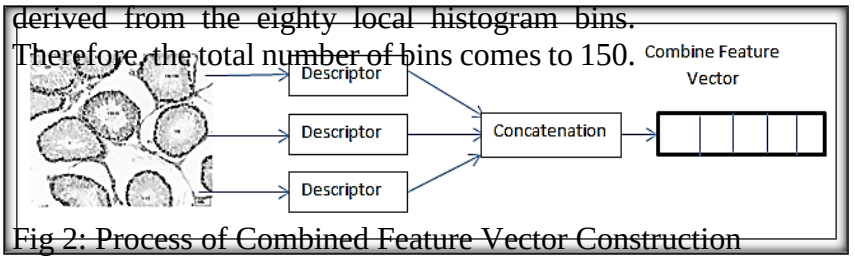

Fig 2: Process of Combined Feature Vector Construction

3.4 Feature Vector Construction

In machine learning and computer vision, the feature vector is used to characterize some object, feature vector consists of n-dimensions each dimension of the vector contains some numerical information regarding an object, which is used to encourage statistical investigation. Fig 2 shows the process of feature vector construction.

3.5 Classification

We have utilized support vector machine (SVM), which is widely adopted super-vised learning model [29]. The studies show that SVM with an early fusion of several features produces the best performance. For multiclass SVM, we have considered chi-square kernel in the form of Equation 3, here are dimensional inputs and K is the kernel function that maps the data from dimensional space to dimension space, usually is much larger than. Studies show that if features are in the form of a histogram, the Chi-square kernel yields better performance as compared to the other kernels 30.

$$K(X,Y) = \sum_i x_i \frac{y_i}{x_i + y_i} \quad (3)$$

For the purposes of categorization in this study project, we made use of the LIB-SVM toolkit. We used the one-vs-all technique, and as part of this approach, we trained a binary classifier that was distinct to each class. This allowed us to handle the difficulty of multi-class classification. Examples that are considered to be in a favorable light are those that correspond to the aforementioned category, whilst situations that have not yet been settled are erred to as examples that are considered to be in a negative light. There were numerous substantial irregularities within the data set. A few classes have an extremely low number of training instances, and while working with them, binary classification learners experience unbalanced distributions even though the class distribution is balanced in the training set. This is because the set of negatives people observe is often a lot larger than the set of positives they see, and this is the reason why this is the case. Because of this, the model almost always considers the one-vs-all method to be lopsided in terms of power distribution. In order to find a solution to this issue, we gave each positive and negative class a certain amount of weight. [25].

4 Evaluation and Results Details

The performance of proposed framework is evaluated in the terms of Accuracy, Precision, Recall, F-score and ROC analysis. In following equations and represents true positive, false positive, true negative, false negative respectively.

$$Accuracy = \frac{TP + TN}{T + N} \quad (4)$$

$$Precision = \frac{TP}{(TP + FP)} \quad (5)$$

$$Recall = \frac{TP}{(TP + FN)} \quad (6)$$

$$F = \frac{2(Precision.Recall)}{(Precision + Recall)} \quad (7)$$

4.1 Dataset

We have tested the performance of our proposed model using the 2D HeLa Data set [6]. Dataset downloaded from link http://murphylab.web.cmu.edu/data/. This dataset includes 862 different fluorescence microscopy images that have been categorized into a total of ten distinct groups. In order to provide the simulation with even more accurate results, the CHO dataset that was developed in the AIIA lab was also utilized [26]. The dataset in question will be obtained via the aforementioned website. https://ome.grc.nia.nih.gov/iicbu2008/hela/index.html. The details of both the data sets are presented in Tables 1 and 2, respectively.

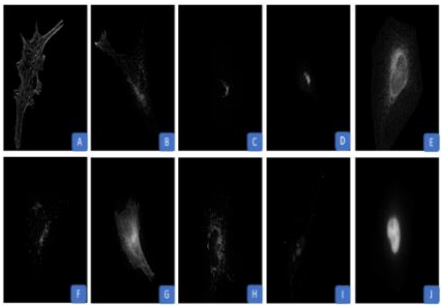

Fig 3: Representative images from the classes of 2D Hela dataset used as input (A) ActinFilaments, (B) Endosomes, (C) ER, (D) Golgi_gia, (E) Golgi_gpp, (F) Lysosome, (G) Microtubules, (H) Mitochondria, (I) Nucleolus, (J) Nucleus.

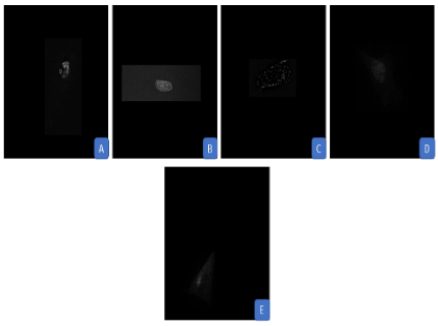

Fig 4: Representative images from the classes of CHO dataset used as input (A) giantin, (B) Hoechst, (C) lamp2, (D) nop4, (E) tubulin

Table 1 2D Hela Dataset

	Class Name									
	ActinFilaments	Nucleus	Endosomes	ER	Golgi Giantin	Golgi GPP130	Lysosome	Microtubules	Mitochondria	Nucleo
curacy	98	87	91	86	87	85	84	91	73	80

Table 2 CHO Dataset

	Class Name				
	Giantin	Hoechst	Lamp2	Nop4	Tubulin
Accuracy	77	67	97	33	51

4.2 Results for Individual Feature Extraction Methods

4.2.1 Classification on Local Binary Pattern Descriptor (LBP)

We extracted LBP feature from the image based on different size radius and a different number of neighbors. Table 3 and 4 shows the results of experiments on Hella dataset and CHO dataset. Many image regions are relatively uniform, and it is valid to investigate whether the robustness of the features can be improved in these regions. That is why we did not consider LBP alone for classification.

Table 4 Class Wise Accuracy Using LBP on CHO Dataset

	Class Name									
	ActinFilaments	Nucleus	Endosomes	ER	Golgi Giantin	Golgi GPP130	Lysosome	Microtubules	Mitochondria	Nucl
racy	100%	93.10%	73.62%	66.27%	95.40%	63.52%	88.09%	89.01%	65.75%	98.7

	Class Name				
	Giantin	Hoechst	Lamp2	Nop4	Tubulin
Accuracy	96.10%	0%	100%	100%	100%

1) **Classification Based on Edge Histogram Descriptor**

EHD characterizes local edge spreading in the image; the features are computed for the whole image by setting the threshold value to 0.10. Table 5 and 6 shows the obtained results for edge histogram descriptor on Hella dataset and CHO dataset.

Table 5 Average Accuracy Using EHD on 2D Hela Dataset

	Class Name									
	ActinFilaments	Nucleus	Endosomes	ER	Golgi Giantin	Golgi GPP130	Lysosome	Microtubules	Mitochondria	Nucl
racy	100%	95.40%	93.40%	98.83%	97.70%	95.29%	95.23%	97.80%	90.41%	98.7

	Class Name				
	Giantin	Hoechst	Lamp2	Nop4	Tubulin
Accuracy	97.40%	100%	100%	100%	100%

4.2.2 Classification Based on SFTA Features

SFTA (Segmentation-based Fractal Texture Analysis) extracts texture features from an image. Table 7 and 8 shows the obtained results for SFTA on Hella dataset and CHO dataset.

Table 7 Classification Performance of SETA on 2D Hela Dataset

	Class Name									
	ActinFilaments	Nucleus	Endosomes	ER	Golgi Giantin	Golgi GPP130	Lysosome	Microtubules	Mitochondria	Nucleolus
Accuracy	97.95%	95.40%	71.42%	81.39%	97.70%	85.88%	80.95%	84.61%	57.53	93.75%

Table 8 Classification Performance of SETA on CHO Dataset

	Class Name				
	Giantin	Hoechst	Lamp2	Nop4	Tubulin
Accuracy	86.61%	98.55%	93.81%	100%	94.11%

4.3 Results for Hybrid Features

Several studies have shown that use of diverse features captures different infor-mation of the image and thus, their combination offers a complete representation of the visual contents of an image and clearly provides better performance as compared to single feature approach [11, 25]. Classification results are presented in the form of accuracy, precision, recall and F-score. Table 9 displays the detail performance in the term of accuracy for every individual class for 2D Hela dataset.

4.3.1

Class Name	**Accuracy**	**Precision**	**Recall**	**F1-Score**	**Area Under Curve**
Actin	100	1	1	1	1.0000
Endosome	98.8372	0.9884	1	0.9942	0.9904
ER	95.8044	0.9560	0.9886	0.9721	0.9895
Golgi_gia	100	1	0.9775	0.9886	1.0000
Golgi_gpp	97.6471	0.9765	0.9881	0.9822	0.9895
Lysosome	97.6190	0.9762	0.9862	0.9762	0.9851
Microtubules	100	1	1	1	1.0000
Mitochondria	100	1	0.9605	0.9799	0.9962
Nucleolus	100	1	1	1	1.0000
Nucleus	100	1	1	0.59	1.0000

ROC Analysis

ROC analysis is also carried out on results of hybrid features. The area under the ROC curve is calculated for each class and then average area under the curve is also computed. Fig 5 shows the ROC plot for all 10 classes.

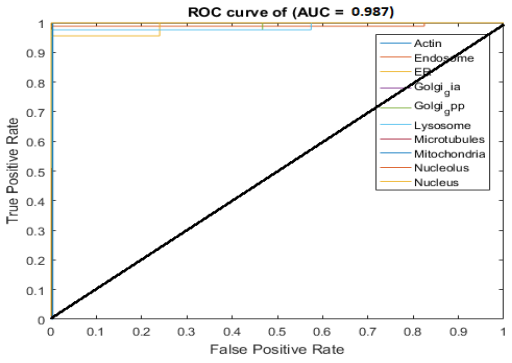

Fig 5 ROC Curve for 2D Hela Dataset using SVM classification

4.4 Performance Comparison

The best accuracy reported by using 2D Hela cells dataset is 98% and by using CHO dataset is 99%. Obtained accuracy by proposed methodology is 98.9% for Hella Cells dataset and 99.6% for CHO dataset, which is 0.9 % and 0.6 % higher than [1,12] respectively.

According to best of our knowledge, the conclusion from experimentations is that the results acquired with our methodology are better from previously reported results. It is due to fact that the dataset is so much diverse that single feature extraction algorithm with particular settings is not sufficient to extract information that provides better classification and retrieval results. We utilized a different type of features ex-traction algorithms that capture texture and other visual information; furthermore, we tried to employ different setting for every algorithm to extract more and more information that helped us in the classification task.

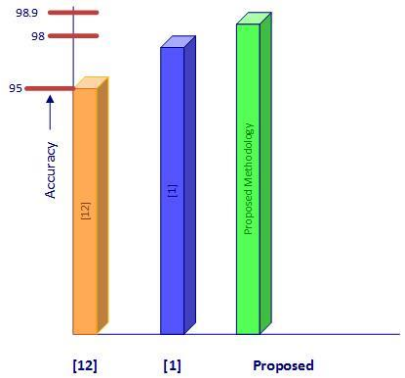

Figure 1Fig 6 Performance Comparison

The yellow bar graph shows the accuracy reported by using 2D Hela cells dataset [12] is 95% and yellow bar graph shows the accuracy in [1] is 98% and green bar shows accuracy achieved by our proposed framework is 98.9%. Obtained accuracy by proposed methodology is 0.9% higher than [1,12] respectively.

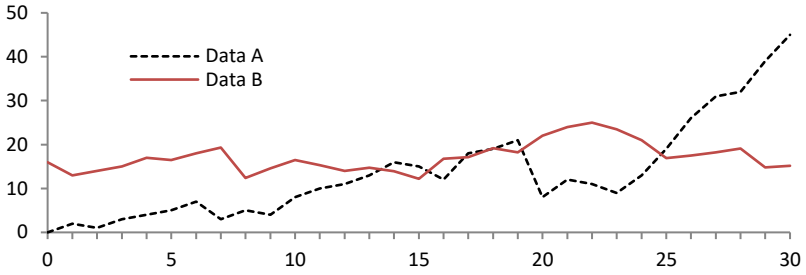

Fig. 7. A figure caption is always placed below the illustration. Short captions are centered, while long ones are justified. The macro button chooses the correct format automatically

.

5 Conclusion

In this research, we present the methodology in detail that we utilized for the fluorescent microscope images. More particularly, we utilized the dataset of the 2D Hela cells and CHO. 2D Hela cells dataset contains 10 classes and CHO dataset con-tains 5 classes. We experimented with several types of visual features, the features that we employed in our research consist of, Haralick's Features, SFTA Features, local binary pattern (LBP), edge histogram descriptor (EHD). First, we experimented with each individual feature extraction technique. We also experimented with com-bined visual features obtained by different settings for SFTA, Haralick's, LBP, and EHD. Finally, the experiments are carried out for the hybrid feature. The experimen-tation (for individual feature extraction method) reveals that the SFTA are best per-forming individual features in our framework and exhibited the accuracy of 90% and 96% on 2D Hela and CHO dataset respectively. Haralick's features demonstrated second best performance as individual feature extraction techniques and exhibited the accuracy of 73% and 72% respectively on 2D Hela and CHO dataset. Other fea-tures like color edge histogram descript (EHD) exhibited the accuracy of 63%, 88 % respectively on 2D Hela and CHO dataset.

After selection of best performing set of features, early fusion (Concatenation) is applied on features, to form a combined hybrid feature vector and experiments car-ried out on hybrid feature vector. The hybrid features exhibited the

accuracy of 98.9% on 2D Hela dataset and 99.6% on CHO dataset. The comparison of the per-formance for hybrid features is compared with other performances mentioned in the literature, on this dataset the best-reported accuracy in 2D Hela dataset was 98%. The results of our proposed methodology are approximately 0.9% and 0.6% higher than the best reported results on the 2D Hela dataset and the CHO dataset, respectively. This is because the use of diverse features captures different information of image, and as a result, their combination offers a complete representation of the visual content of an image. This is due to the fact that, the use diverse features cap-ture different information of image, and as a result, their combination offers a complete representation of the visual content of an image.

Data Availability Statement

Data are available from the authors upon reasonable request and with permission of Author.

References

[1] Boland, M.V. and Murphy, R.F., 2001. A neural network classifier capable of recognizing the patterns of all major subcellular structures in fluorescence microscope images of HeLa cells. Bioinformatics, 17(12), pp.1213-1223.

[2] Khan, S.U.R.; Zhao, M.; Asif, S.; Chen, X.; Zhu, Y. GLNET: Global–local CNN's-based informed model for detection of breast cancer categories from histopathological slides. J. Super-comput. 2023, 80, 7316–7348.

[3] Murphy, R.F., Boland, M.V. and Velliste, M., 2000, August. Towards a Systematics for Protein Subcellular Location: Quantitative Description of Protein Localization Patterns and Au-tomated Analysis of Fluorescence Microscope Images. In ISMB (Vol. 8, pp. 251-259).

[4] Khan, S.U.R.; Zhao, M.; Asif, S.; Chen, X. Hybrid-NET: A fusion of DenseNet169 and advanced machine learning classifiers for enhanced brain tumor diagnosis. Int. J. Imaging Syst. Technol. 2024, 34, e22975.

[5] Khan, S.U.R.; Asif, S.; Bilal, O.; Ali, S. Deep hybrid model for Mpox disease diagnosis from skin lesion images. Int. J. Imaging Syst. Technol. 2024, 34, e23044.

[6] Khan, S.U.R.; Raza, A.;Waqas, M.; Zia, M.A.R. Efficient and Accurate Image Classification Via Spatial Pyramid Matching and SURF Sparse Coding. Lahore Garrison Univ. Res. J. Com-put. Sci. Inf. Technol. 2023, 7, 10–23.

[7] Chen, S.C. and Murphy, R.F., 2006. A graphical model approach to automated classification of protein subcellular location patterns in multi-cell images. BMC bioinformatics, 7(1), p.90.

[8] Raza, A.; Meeran, M.T.; Bilhaj, U. Enhancing Breast Cancer Detection through Thermal Imaging and Customized 2D CNN Classifiers. VFAST Trans. Softw. Eng. 2023, 11, 80–92.

[9] Dai, Q., Ishfaque, M., Khan, S. U. R., Luo, Y. L., Lei, Y., Zhang, B., & Zhou, W. (2024). Image classification for sub-surface crack identification in concrete dam based on borehole CCTV images using deep dense hybrid model. Stochastic Environmental Research and Risk Assess-ment, 1-18.

[10] Khan, S. U. R., & Asif, S. (2024). Oral cancer detection using feature-level fusion and novel self-attention mechanisms. Biomedical Signal Processing and Control, 95, 106437.

[11] Khan, U. S., & Khan, S. U. R. (2024). Boost diagnostic performance in retinal disease clas-sification utilizing deep ensemble classifiers based on OCT. Multimedia Tools and Applications, 1-21.

[12] Tahir, M. and Khan, A., 2016. Protein subcellular localization of fluorescence microscopy images: employing new statistical and Texton based image features and SVM based ensemble classification. Information Sciences, 345, pp.65-80.

[13] Farooq, M. U., Khan, S. U. R., & Beg, M. O. (2019, November). Melta: A method level energy estimation technique for android development. In 2019 International Conference on Inno-vative Computing (ICIC) (pp. 1-10). IEEE.

[14] Shahzad, I., Khan, S. U. R., Waseem, A., Abideen, Z. U., & Liu, J. (2024). Enhancing ASD classification through hybrid attention-based

learning of facial features. Signal, Image and Video Processing, 1-14.
[15] Lin, C.C., Tsai, Y.S., Lin, Y.S., Chiu, T.Y., Hsiung, C.C., Lee, M.I., Simpson, J.C. and Hsu, C.N., 2007. Boosting multiclass learning with repeating codes and weak detectors for pro-tein subcellular localization. Bioinformatics, 23(24), pp.3374-3381.
[16] HUSSAIN, S., RAZA, A., MEERAN, M. T., IJAZ, H. M., & JAMALI, S. (2020). Domain Ontology Based Similarity and Analysis in Higher Education. IEEEP New Horizons Journal, 102(1), 11-16.
[17] Khan, M. A., Khan, S. U. R., Haider, S. Z. Q., Khan, S. A., & Bilal, O. (2024). Evolving knowledge representation learning with the dynamic asymmetric embedding model. Evolving Systems, 1-16.
[18] Lin, Y.S., Lin, C.C., Tsai, Y.S., Ku, T.C., Huang, Y.H. and Hsu, C.N., 2010. A spectral graph theoretic approach to quantification and calibration of collective morphological differences in cell images. Bioinformatics, 26(12), pp. i29-i37.
[19] Mei, S., 2012. Multi-kernel transfer learning based on Chou's PseAAC formulation for protein submitochondria localization. Journal of Theoretical Biology, 293, pp.121-130.
[20] Shi, J.Y., Zhang, S.W., Pan, Q., Cheng, Y.M. and Xie, J., 2007. Prediction of protein sub-cellular localization by support vector machines using multi-scale energy and pseudo amino acid composition. Amino acids, 33(1), pp.69-74.
[21] Khan, S. U. R., Zhao, M., Asif, S., Chen, X., & Zhu, Y. (2023). GLNET: global–local CNN's-based informed model for detection of breast cancer categories from histopathological slides. The Journal of Supercomputing, 1-33.
[22] Wajid, M., Abid, M. K., Raza, A. A., Haroon, M., & Mudasar, A. Q. (2024). Flood Predic-tion System Using IOT & Artificial Neural Network. VFAST Transactions on Software Engi-neering, 12(1), 210-224.
[23] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 7, pp. 971–987, Jul. 2002.
[25] Raza, A., Soomro, M. H., Shahzad, I., & Batool, S. (2024). Abstractive Text Summarization for Urdu Language. Journal of Computing & Biomedical Informatics, 7(02).
[26] Al-Khasawneh, M. A., Raza, A., Khan, S. U. R., & Khan, Z. (2024). Stock Market Trend Prediction Using Deep Learning Approach. Computational Economics, 1-32.
[27] Meeran, M. T., Raza, A., & Din, M. (2018). Advancement in GSM Network to Access Cloud Services. Pakistan Journal of Engineering, Technology & Science [ISSN: 2224-2333], 7(1).
[28] Khan, U. S., Ishfaque, M., Khan, S. U. R., Xu, F., Chen, L., & Lei, Y. (2024). Comparative analysis of twelve transfer learning models for the prediction and crack detection in concrete dams, based on borehole images. Frontiers of Structural and Civil Engineering, 1-17.
[29] J. M. Keller, M. R. Gray, and J. A. Givens, "A fuzzy k-nearest neighbor algorithm," IEEE Transactions on Systems, Man, and Cybernetics, vol. SMC-15, no. 4, pp. 580–585, Jul. 1985.
[30] A. Vedaldi and A. Zisserman, "Efficient additive kernels via explicit feature maps," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 34, no. 3, pp. 480–492, Mar. 2012.
[31] Farooq, M. U., & Beg, M. O. (2019, November). Bigdata analysis of stack overflow for energy consumption of android framework. In 2019 International Conference on Innovative Computing (ICIC) (pp. 1-9). IEEE.
[32] von F, Z., 1934. Beugungstheorie des schneidenver-fahrens und seiner verbesserten form, der phasenkontrastmethode. Physica, 1(7-12), pp.689-704.
[33] Boland, M.V. and Murphy, R.F., 1998. After sequencing: quantitative analysis of protein localization. IEEE engineering in medicine and biology magazine: the quarterly magazine of the Engineering in Medicine & Biology Society, 18(5), pp.115-119.
[34] Jurvelin, J.S., Müller, D.J., Wong, M., Studer, D., Engel, A. and Hunziker, E.B., 1996. Surface and subsurface morphology of bovine humeral articular cartilage as assessed by atomic

force and transmission electron microscopy. Journal of structural biology, 117(1), pp.45-54.
[35] Jennrich, R.I. and Sampson, P., 1977. Stepwise discriminant analysis. Statistical methods for digital computers, 3, pp.77-95.
[36] Huang, K., Lin, J., Gajnak, J.A. and Murphy, R.F., 2002. Image content-based retrieval and automated interpretation of fluorescence microscope images via the protein subcellular location image database. In Biomedical Imaging, 2002. Proceedings. 2002 IEEE International Symposi-um on (pp. 325-328). IEEE.
[37] Roques, E.J. and Murphy, R.F., 2002. Objective evaluation of differences in protein subcel-lular distribution. Traffic, 3(1), pp.61-65.
[38] Murphy, R.F., Velliste, M. and Porreca, G., 2003. Robust numerical features for description and classification of subcellular location patterns in fluorescence microscope images. Journal of VLSI signal processing systems for signal, image and video technology, 35(3), pp.311-321.
[39] Raza, A., & Meeran, M. T. (2019). Routine of Encryption in Cognitive Radio Network. Mehran University Research Journal of Engineering and Technology [p-ISSN: 0254-7821, e-ISSN: 2413-7219], 38(3), 609-618
[40] Huang, K. and Murphy, R.F., 2004, April. Automated classification of subcellular patterns in multicell images without segmentation into single cells. In Biomedical Imaging: Nano to Mac-ro, 2004. IEEE International Symposium on (pp. 1139-1142). IEEE.
[41] Hu, Y., Osuna-Highley, E., Hua, J., Nowicki, T.S., Stolz, R., McKayle, C. and Murphy, R.F., 2010. Automated analysis of protein subcellular location in time series images. Bioinformat-ics, 26(13), pp.1630-1636.
[42] Huh, S., Lee, D. and Murphy, R.F., 2009. Efficient framework for automated classification of subcellular patterns in budding yeast. Cytometry Part A, 75(11), pp.934-940.
[43] Hyvärinen, A., Hurri, J. and Hoyer, P.O., 2009. Natural Image Statistics: A Probabilistic Approach to Early Computational Vision (Vol. 39). Springer Science & Business Media.
[44] De Carvalho, M.A., Lotufo, R.D.A. and Couprie, M., 2007. Morphological segmentation of yeast by image analysis. Image and Vision Computing, 25(1), pp.34-39.
[45] Hyvärinen, A., Hurri, J. and Hoyer, P.O., 2009. Natural Image Statistics: A Probabilistic Approach to Early Computational Vision (Vol. 39). Springer Science & Business Media.
[46] Zhao, T., Velliste, M., Boland, M.V. and Murphy, R.F., 2005. Object type recognition for automated analysis of protein subcellular location. IEEE transactions on image processing, 14(9), pp.1351-1359.
[47] Danckaert, A., Gonzalez-Couto, E., Bollondi, L., Thompson, N. and Hayes, B., 2002. Au-tomated recognition of intracellular organelles in confocal microscope images. Traffic, 3(1), pp.66-73.