An Evolutionary Approach for Determining Hidden Markov Model for Medical Image Analysis

J. Goh¹, H. L. Tang¹

T. Peto², and G. Saleh³

¹ Department of Computing, University of Surrey,
Guildford, Surrey, GU2, 7XH, UK.
j.goh@surrey.ac.uk

²Reading Centre, Department of Research and
Development, Moorfields Eye Hospital NHS
Foundation Trust, 162 City Road, London,
EC1V 2PD, UK.

³Moorfields Eye Hospital, South Wing, Bedford,
MK42 9DJ, UK

Abstract—Hidden Markov Model (HMM) is a technique highly capable of modelling the structure of an observation sequence. In this paper, HMM is used to provide the contextual information for detecting clinical signs present in diabetic retinopathy screen images. However, there is a need to determine a feature set that best represents the complexity of the data as well as determine an optimal HMM. This paper addresses these problems by automatically selecting the best feature set while evolving the structure and obtaining the parameters of a Hidden Markov Model. This novel algorithm not only selects the best feature set, but also identifies the topology of the HMM, the optimal number of states, as well as the initial transition probabilities.

Keywords—component; Hidden Markov Models; Memetic Algorithms; Genetic Algorithms; Particle Swarm Optimisation; Contextual Reasoning, Diabetic Retinopathy

I. INTRODUCTION

Diabetic Retinopathy (DR) is a complication of diabetes and if left untreated, vision loss is unavoidable. Micro aneurysms are one of the first visible signs of DR and it is known that quantities of this clinical sign can help diagnose the progression of the disease. Micro aneurysms are swelling of the capillaries that are caused by the weakening of the vessel walls due to high sugar levels in diabetes and eventually leak to produce exudates. In retina images, micro aneurysms appear as small reddish dots with similar intensity as haemorrhages and blood vessels. This particular sign is an important early indicator of the disease and can contribute to helping ophthalmologists identify effective treatment for the patient at an early stage.

The problem of scalable image recognition has long been a research issue. Accurate detection of micro aneurysms over large-scale data is a challenging task which have been the focus of many researchers for over the last decade [1, 2, 3, 4]. The main obstacles are due to the variability in the retina image caused by multiple factors such as degree of pigmentation of epithelium and choroid in the eye, size of pupil, illumination, disease, imaging settings (which can vary

even with same equipment), patients' ethnic origin, and other variants.

Researchers have used standard image processing and classification techniques to deal with the ambiguity in detecting micro aneurysms (MA). Niemeijer et al. [2] and Sinthanayothin et al. [3] adopted image processing techniques to first extract useful features followed by recognition through a classifier. However, micro aneurysms are often mistaken as other similar visual content such as fine ends of the blood vessels or noise resulting in false positives. Hence, appropriate contextual constraints should be sought to minimize the ambiguity in the image analysis process. The question is which features should be used so as to capture the variation within the data set. When the scale of the application is large, (our ongoing project is currently dealing with overall 20,000 images with an ultimate goal of a consistent performance on millions of patients' images), it is more complex to determine suitable features that represent largely varied visual content. Adopting more features may cover variability but will be inevitably irrelevant or redundant for some situation apart from having high dimensional problems causing prediction errors. Thus, feature selection is seen as an important aspect of this research.

A Hidden Markov Model (HMM) is a statistical modelling approach and is highly capable of modelling the structure of an observation sequence. More importantly, they are capable of modelling the context dependent entities by allowing fine details to be learnt through the data by adjusting the transition probabilities and emission probabilities. While HMMs have been successful in many applications such as speech recognition [5, 6, 7], DNA sequencing [8] and handwriting recognition [9], very little research has been performed to statistically model and understand the high level context within images beyond pixel level. In speech recognition, HMMs can determine the statistical variations of utterance from occurrence to occurrence. Similar principle can be

applied to understand the different context in which the feature appears within the image. In this work, HMM is to encapsulate the contextual information within the underlying sequential states to model the contextual information of features in diabetic retinopathy screening images.

In this paper, HMMs is used to model the context in which micro aneurysms appear. Traditionally, training a HMM can usually be carried out through the refinement of the HMM after each training. Such refinement includes changing the number of states, the initial distribution states and the transition probabilities before re-training the HMM. The most popular training algorithm for HMM is the Baum-Welch [10] algorithm; however, this algorithm is a hill climbing algorithm and heavily depends on the initiate estimates. While the use of HMM has proved to be generic in nature, there are a few issues surrounding the success of a HMM. Firstly, how to determine the topology of the HMM and secondly, what are the optimal model parameters for the model to best represent the observation. Lastly, the training of HMM is computationally intensive and there is no known method that can guarantee to obtain the most optimal model.

Kwong et al. [11] and Bhuriyakorn et al. [12] have applied Genetic Algorithms to evolve the structure of the HMM for speech recognition. In similar work, Kyoung-Jae et al. [8] applied this technique to DNA recognition. However, in their research, only the optimal number of states and the structure of the HMM were obtained. No consideration was given to the transition probabilities of the HMM which are crucial during the re-estimating of the model in training process using Baum Welch algorithm, which heavily relies on good estimates. In other words, obtaining an optimal Hidden Markov Model requires a balance between a good topology and good initial parameters. Furthermore, most of these techniques use manually selected features. Automatic selection of features can potentially offer a better set of (combined) features which may not be easily observed manually but better suit the complex data. This is demonstrated in this work by a comparison between these two approaches.

Our previous work [13] has demonstrated the effectiveness of HMM in the detection of micro aneurysms as a contextual analysis model. This work is to further develop this method by integrating an automated feature selection process so as to enable the context model to maximise and generalise its understanding of the ground truth in the training data. We propose the use of Memetic algorithms to determine a suitable feature set whilst optimising the Hidden Markov Model. In Section 2, we give a brief description on the algorithm used. The technique used for optimising the HMM is presented in Section 3. Section 4 discuss the experimental results and we summarise our work in section 5.

II. HYBRID EVOLUTIONARY OPTIMISATION & HIDDEN MARKOV MODELS

Memetic algorithms are a class of hybrid searching

algorithms that involves the integration of population-based approach and local improvement. Population based Genetic Algorithms have known to be slow in convergence because of their failure to exploit the information in the local regions. Memetic algorithms combine a population based global heuristic search strategy with a local refinement [14]. Such Memetic algorithms aim to achieve a balance between the exploration and the exploitation of the search space to maximise the search performance and have seen success in many areas.

The Memetic algorithm adopted in this research combines a population-based global heuristic search strategy to determine an optimum feature set for HMM as well as the structure of the HMM, and a Particle Swarm Optimization to obtain an optimised initial transition probability for a fully optimized HMM. The following sections provide details of this work.

A. Genetic Algorithms

Genetic algorithms (GA) are a type of evolutionary algorithms and are known for their explorative capabilities in problem spaces. GA is chosen as the mechanism used to evolve the feature sets and the structure of the HMM. At each generation of the GA, a new set of solutions is created by a process of selecting individuals according to their strengths (fitness) in the problem domain and genetically modifying them to produce offspring. This process leads to the generation of a new population of individuals that are better suited for the problem than the individuals that they are created from, eventually reaching an optimal solution.

B. Particle Swarm Optimisation

Particle Swarm Optimisation (PSO) is another evolutionary algorithm that differs from other population based optimisation techniques. Unlike GA that manipulates the individual, PSO relies on a one way sharing mechanism by using only the best particle to give out information as compared to GA that shares information among themselves. This is a parallel search algorithm that functions by propelling the individual solution (particle) through the search space with a velocity that is dynamically modified based on its own strength and the strength of other particles in the swarm. Hence, allowing PSO the capabilities to converge faster to an optimal solution [15].

III. MEMETIC ALGORITHM TRAINING

In order to illustrate the use of evolutionary algorithms in this research, sub regions of particular diabetic retinopathy (DR) signs are represented using optimised Hidden Markov Models evolved through evolutionary computation. This integrated solution evidently not only automates the discovery of the appropriate feature set for DR signs, but also the HMM as well. The resulting model can also attain a better accuracy while avoiding over-fitting.

In this research, the Memetic algorithms consists of two Genetic Algorithms and one PSO with one GA to determine the feature set and the other for evolving the structure of the HMM, while using PSO to optimize the Transition Probability Matrix. Figure 1 illustrates the pseudo code for this algorithm. Details of the key functions are further discussed in the next section.

```

Initialise Population
While iteration < Max_Generation
  -- Commence GA for Feature Selection --
  Selection of a set of solutions using Roulette Wheel Selection;
  Crossover: only on the visual feature part of chromosome;
  Calculate Fitness of all solutions;
  -- Commence GA for HMM --
  For all the solutions in the population
    Selection of a set of solutions using Roulette Wheel Selection;
    Crossover: only on HMM state part of chromosome;
    Re-calculate Fitness for all solutions;
    Rank solutions
  -- Commence PSO for HMM Transition Probability Generation --
  For the top 20% of the fittest solutions in the population
    Generate Transition Probability Matrix using PSO
    Re- calculate Fitness for the solution
    If new fitness value > previous fitness value
      Replace current transition probability with the new one
    endIf
  endFor
endFor -- end PSO --
Performance Recombination to replace weaker solutions with fitter
newly generated solutions
-- End GA HMM --
-- End GA for Feature Selection --
endWhile

```

Figure 1. Pseudocode for Memetic Algorithm

The GA will first perform the exhaustive search to determine the visual features for each solution. Each generated solution will then be passed to another GA which will evolve the structure of the HMM. PSO will be carried out on the fitter solutions to further optimise them. Ideally, when the termination criteria have been met, the final population would consist only of the best individuals which would be decoded as the optimised set of solutions.

A. Genetic Algorithm for Feature Selection and HMM structure

In this study, DR features such as micro aneurysms are modelled using HMM directly from the visual features rather than from classifiers. The candidate 33x33 pixels image region is divided into 9 smaller sub images as illustrated in Figure 2. Image feature extraction is applied to each of these 11x11 smaller sub images and the selected features (as listed in Table I) are used as a sequence of observations for the HMM.

Figure 2. Example of Micro aneurysms Sub Image

Based on a set of visual features, a chromosome will contain elements including *visual features*, *HMM states and transition probabilities*. Visual features are those types listed in Table I but represented by their index numbers in the chromosome. In a chromosome, there are also elements representing an ordered set of HMM state transition types as shown in Table II. The number of states as well as a transition probability matrix corresponding to the ordered set of state transition types are also included in the chromosome.

In HMM, there can be a few different kinds of transitions to be assigned to each state as listed in Table II. These transitions are randomly assigned to each state and encoded into the chromosome using its corresponding transition type code. Initial state transition probabilities are also randomly assigned between the initiating states and the transiting states. The training algorithm used is the Baum Welch (BW) algorithm. It performs the training by using the transition probabilities matrix as the parameters of the HMM. As the training procedure converges, it will adjust the parameters of the HMM accordingly so as to increase the probability of the model assigned to the training set. This is the reason why BW algorithm relies heavily on the initial transition probabilities.

1) Fitness Evaluation

In order to measure the generalisation capability of the HMM for recognising sub-images of micro aneurysm, a fitness evaluation mechanism to gauge the confidence level of each solution is used. The fitness evaluation throughout the whole evolution in this experiment is the same and is defined as the accuracy of the solution over an unseen image test data set,

$$f_v(k) = \frac{\text{Number of Correctly Recognised Test Samples by HMM}(k)}{\text{Total Number of Test Samples}} \quad (1)$$

where k is the k^{th} solution, $HMM(k)$ is the HMM based on the chromosome content of k^{th} solution including visual features and HMM parameters.

TABLE I. LIST OF POSSIBLE FEATURES

Chromosome code	Feature
1	Discrete Cosine Transform
2	Standard Deviation of Red Component
3	Standard Deviation Green Component
4	Standard Deviation Blue Component
5	Standard Deviation L Component of LuV
6	Standard Deviation U Component of LuV
7	Standard Deviation V Component of LuV
8	Standard Deviation H Component of HSI
9	Standard Deviation S Component of HSI
10	Standard Deviation I Component of HSI
11	Third Moment – Measure Histogram
12	Uniformity
13	Entropy – Measures Randomness

TABLE II. POSSIBLE STATE TRANSITION TYPES

Transitions	Models
Type 1
Type 2
Type 3
Type 4
Type 5

2) Selection

Selection is the phase used to determine which parents to choose for reproduction. Similarly, the selection algorithm used for selection parents is the same throughout this experiment and in this work, the Roulette Wheel Selection (RWS) is chosen as the selection algorithm. Each solution in the population will be assigned a probability of selection based on the fitness value it achieved. If solution k has a fitness value as in equation 1, the probability is calculated using:

$$probability(k) = \frac{fv(k)}{\sum_{j=1}^n fv(j)} \quad (2)$$

where, n is the total number of solutions in the population. This procedure will normalise the fitness values and assign probabilities to each solution on the Roulette Wheel.

3) Crossover

1-point crossover is used throughout this experiment. The only difference for feature selection and HMM structure optimisation is the reference table used during the decoding stage to train the HMM. For example, if crossover is performed during feature selection, the corresponding codes in Table I will be used otherwise the content in Table II will be used. In either case, during crossover, if both parents have the same number of features/states, the creation of offspring is straightforward. However, if the two parents have different number of features/states, there must be a decision on how many features/states the offspring will have. For simplicity, it is assumed that the offspring shall have the average number of features/states between the two parents. To make up for the additional features/states, the offspring will adopt the additional features/states from the parent as illustrated in Figure 3.

Figure 3. Crossover Operator

4) Recombination

Recombination is the operation used to introduce the newly created solutions back into the population. For this experiment, new solutions with better fitness values simply gets introduced back into the population by replacing the weaker solutions.

B. Particle Swarm Optimisation - Local Search

As the Baum Welch algorithm is very sensitive to the initial model parameter, there is a need to obtain a balance between a good topology and good initial model parameters. PSO is applied on each solution passed from the GA algorithm in the previous stage to obtain better solutions. PSO is used to search for a parameter that alters the transition probability matrix for a better HMM. The global search for such parameter can be defined by:

$$f(x) = \max\{f(x_i)\}, \text{ where } x \in \{x_1, x_2, x_3, \dots, x_n\} \quad (3)$$

where x_i is the parameter used to alter the transition probabilities of the HMM. The aim of the function is to find x , such that the function $f(x)$ is the maximum in the search space.

The two main components of the PSO are position, $p_i = (p_1, p_2, \dots, p_n)$ and the velocity, $v_i = (v_1, v_2, \dots, v_n)$ of the particles, where the positions of the particles are calculated using:

$$x_i = x_i + v_i \quad (4)$$

where x_i is the current position of the particle i and v_i is defined as the velocity of particle i as represented using :

$$v_i = v_i + r_1(p_{best} - x_i) + r_2(g_{best} - x_i) \quad (5)$$

with g_{best} being the global best position of the swarm and p_{best} as the particle's best location it has encountered. While r_1 and r_2 are random variables generated between 0 and 1 for each particle.

In this algorithm, each individual particle, p , has a current position in search space x_i , a velocity, v_i , and a personal best position in search space defined as p_{best} . Depending on the optimising function, the personal best position corresponds to the position in search space where particle p_i had the best value as determined by the objective function. As we are looking at finding the maximum accuracy; the best value is again determined by the highest attained accuracy as in equation (1). In addition, the position yielding the highest value amongst all the personal best is called the global best position.

At the end of the operation after 30 iterations, the new transition matrix generated by g_{best} is returned to the chromosome in the GA operation where it overwrites its previous transition matrix. This hybrid procedure will ensure that for every structure of the HMM evolved by the GA, there is an optimised transition probability matrix with an optimal feature set.

IV. EXPERIMENTAL RESULTS

A. Data Set

All the sub images used in this experiment are of size 33 by 33 pixels. These training samples used to train the Hidden Markov Models are obtained from 300 retina images with micro aneurysms in them; all of which were obtained from various sources.

B. Training Samples

1614 background (BG), 1894 micro aneurysms (MA) and 1421 blood vessel (BV) sub images were used to train the model. In order to test the accuracy of the models, another 1000 sub image from each category were used as a test set.

C. Manually Drawn HMM

The manually drawn model consisted of a simple left to right model. According to a paper by Bakis [16], the author concluded that the number of states for the model should be equal to the number of states in the sequences. For this experiment, since the feature set is made out of one 33 by 33 pixel sub image divided into 9 smaller sub images, the feature set used is considered to have 9 sequences. Hence, to test the condition set forth by Bakis [16], the models drawn consist of 7, 8 and 9 states. Since, post analysis involves determining the sub image into the three categories, three different HMM were developed representing micro aneurysms (MA), blood vessels (BV) and the background (BG) category.

The performance of the HMM in this particular experiment is dictated by the accuracy over a test set obtained after the last re-estimation procedure determined by the Baum Welch algorithm. Naturally, to determine which models to use, the trained models with the highest accuracy is used. Table III illustrates the performance of the HMM used to model background, micro aneurysms and blood vessels.

TABLE III. OVERALL ACCURACY FOR MANUALLY DRAWN HMM

Models	Average Maximum Likelihood			
	6 States	7 States	8 States	9 States
MA	45.6%	54.6%	54.89 %	55.05%
BV	44.6%	52.6%	53.6%	52.91%
BG	47.6%	55.6%	54.23%	56.87%

All the manually designed models with 9 states outperformed its counter parts in terms of the accuracy achieved. This emphasizes that the number of states chosen should indeed be identical to the number of states in the observation sequences [16]. In the next section, this hypothesis is further tested with the various Memetic algorithms that were used to evolve the HMM.

D. Memetic HMM (M-HMM)

The Memetic Algorithm was run according to the parameters set in Table IV. In this experiment, due to the expensive executive times, we used relatively limited number of solutions and generations throughout this algorithm. We expect that if more numbers of solutions and generations were used, a better solution should be obtained. Also, usually the performance of the model is gauged through the Average Maximum Likelihood (AML), the higher the AML, the better the model. However, analysis from our previous work [13] shows that this is not necessarily true. This is the reason why the accuracy was used as the fitness evaluation rather than the average maximum likelihood. We also believe that by avoiding continuous training to obtain a better average maximum likelihood, over-fitting can be circumvented.

TABLE IV. PARAMETERS USED FOR MEMETIC HMM

Stages of algorithm	No. of Solutions	Generations
Stage 1: Feature Selection (GA)	5	20
Stage 2: HMM structure (GA)	5	10
Stage 3: Transition Prob. (PSO)	10	15

The results in Table V illustrate the solutions at the different generations of the algorithm. Due to space constrain, the results are only shown at intervals of 5 generations. Table V shows the solutions were increasing steadily through the evolution. This demonstrates that the algorithm used was capable of dynamically selecting the features and optimising the HMM based on the data it was presented with. At the end of the evolution, the best solution converged onto 3 features, which were decoded as the Discrete Cosine Transform, Uniformity and the Entropy of the sub image.

TABLE V. NO. OF FEATURES AND ACCURACY OVER EVOLUTION

Solution	Gen 1 No. of Features /Accuracy	Gen 5 No. of Features /Accuracy	Gen 10 No. of Features /Accuracy	Gen 15 No. of Features /Accuracy	Gen 20 No. of Features /Accuracy
1	2/65.9%	3/67.3%	7/63.5%	6/70.78%	6/70.67%
2	4/55.9%	6/60.7%	5/68.6%	4/78.6%	4/79.6%
3	7/64.2%	6/70.6%	3/69.7%	3/65.5%	4/74.5%
4	8/65.4%	7/66.0%	4/75.3%	3/81.5%	3/81.5%
5	3/66.2%	3/68.5%	5/70.6%	5/71.82%	4/78.5%

We further analyse the results at stage 2 of the algorithm which was the evolution of the HMM structure. At generation 10, Table VI illustrates the evolution of the HMM for solution 3 (*highlighted in Table V and inserted as solution 1 in Table VI*) obtained from the previous stage.

TABLE VI. ACCURACY DURING HMM STRUCTION EVOLUTION AT GENERATION 10

Solution	Gen 1 Number of States/ Accuracy	Gen 5 Number of States/ Accuracy	Gen 10 Number of States/ Accuracy
1	9/69.7%	9/72.5%	9/78.6%
2	8/68.4%	9/73.4%	9/77.1%
3	5/56.8%	6/60.1%	7/68.4%
4	8/65.4%	8/69.2%	9/79.3%
5	7/63.4%	8/68.3%	9/76.5%

Through analysis of the results in Table VI, by the end of the 10th evolution, all but one of the solutions converged onto 9 states which was similar to a manually drawn HMM. This demonstrates that the algorithm is capable of determining the optimal number of states and structure of the HMM. At the end of the generation, this takes us to stage 3 where the best

solution is sent to the PSO for Transition Probability Matrix optimisation. Since solution 4 (*highlighted in Table VI*) from the previous GA HMM evolution obtained the best results, it is used to illustrate the evolution of the Transition Probability Matrix in Table VII.

TABLE VII. THE ACCURACY DURING TRANSITION PROBABILITY MATRIX OPTIMISATION

Generation	Accuracy
Generation 1	79.3%
Generation 3	79.6%
Generation 6	80.2%
Generation 9	81.2%
Generation 12	81.5%
Generation 15	81.5%

Similarly, by the 12 generation, the solution was further enhanced by optimising the Transition Probability Matrix of the HMM structure. The illustrated results suggests that by using this technique, the parameters for each solution are adaptive over the evolutionary process at each stage of the algorithm allowing for the determination of the feature set; the optimised structure of the HMM and its parameters. It also suggests that this technique reduces the risk of over-fitting the training data since the fitness evaluation is the accuracy rather than continuous training for the highest average maximum likelihood that may eventually cause over-fitting.

E. Memetic HMM (No Feature Selection)

We compare the results obtained with our previous work [14], a Memetic HMM but without any form of feature selection (*M-HMM-NF*). The mechanism for this algorithm is similar to the work presented here with the exception of the feature selection module. The feature selected for this particular experiment uses the Discrete Cosine Transform. The experiment was re-run using the same training samples presented earlier.

Comparing M-HMM, M-HMM-NF and the manually drawn HMM in Table VIII, both the evolved HMMs obtained a better accuracy as compared to the manually drawn approach. However, through the automatic determination of the features, the M-HMM obtained better overall accuracy over the M-HMM-NF. The performance listed in Table VI, indicates that the optimal number of states obtained by the M-HMM is also identical to a manually trained HMM with 9 states. Naturally, for each model, the model with the highest accuracy was used.

Although the best solution gave an overall accuracy of only 81.5%, the method presented is highly capable of ensuring a fully optimised HMM. This HMM was developed as part of a micro aneurysms detector that includes a multiple classifier module. Its purpose was to provide post analysis to

filter out the false positives (i.e non MA regions detected erroneously as MA) from the classification module. Figure 4 illustrates the detector's ability to detect micro aneurysms (labelled in black boxes).

TABLE VIII. COMPARISON BETWEEN M-HMM, M-HMM-NF AND MANUALLY DRAWN HMM

Models	Individual Accuracy		
	M-HMM	M-HMM-NF	Manual HMM
MA	83.18%	75.18%	55.05%
BV	77.2%	69.59%	52.91%
BG	83.07%	72.1%	56.87%

Figure 4. Detected Micro aneurysms

V. CONCLUSION

A novel method using Hidden Markov Models to model the context in which micro aneurysms appear was presented in this paper. The M-HMM was designed to separate the evolution of three different entities: the feature set, the HMM structure and the transition probability matrix for the HMM. Its goal was to determine whether it was possible to automatically determine the feature set based on ground truths while evolving the structure of the HMM as well as its transition probability matrix. Using this method, the HMM is able to filter out false positives caused by background noise during pre-processing and fine ends of the blood vessels.

A comparison was made between the two evolved HMM using Memetic Algorithms and the manually drawn HMM that was designed. Our results show that the evolved HMM converged to the same number of states as a manually designed HMM. Furthermore, both the evolved HMM produced better results as well, with the M-HMM performing better than the evolved model with no feature selection. Through analysis of the results at different stages of the evolution, it can be seen that the accuracy improved steadily through the various stages. The hybrid algorithm used to evolve the HMM also proved that contextual information could work together through an evolutionary process for feature selection. Using this technique, the evolved HMM obtained an overall accuracy of 81.5% with an individual accuracy of 83.18%, 77.2% and 83.07% for MA, BV and BV

models respectively, outperforming both the M-HMM-NF and the manually designed HMM.

While the overall accuracy is only 81.5%, the HMM is part of a micro aneurysms detector which is integrated into a larger system. The goal of the system is to separate normal and abnormal DR screening images by obtaining information from multiple processors and dynamically fusing the information at the last stage for final image interpretation. Currently, we are testing the integrated system on diverse populations with each population over 10,000 images.

ACKNOWLEDGEMENT

This work is primarily in collaboration with the Reading Centre, Department of Research and Development, Moorfields Eye Hospital NHS Foundation Trust, United Kingdom. The authors would also like to thank Bedfordshire Retinal Screening Service, United Kingdom, for their invaluable input.

REFERENCES

- [1] Fleming, A. D., Philip, S., Goatman, K. A., Olson, J. A., Sharp, P. F. (2006). Automated microaneurysm detection using local contrast normalization and local vessel detection. *IEEE Trans Med Imaging*, Vol. 25(9), pp. 1223-32.
- [2] Meindert, N., Ginneken, B., Stal, J., Suttrop-Schulten, M., Abramoff, M., 2005. "Automatic Detection of Red Lesions in Digital Color Fundus Photograph", *IEEE Transaction on Medical Imaging*, Vol. 25(5).
- [3] Sinthanayothin, C., Boyce, J. F., Williamsom, T.H., Cool, H.L., Mensah, E., Lai, S., Usher, D., 2002. "Automated Detection of Diabetic Retinopathy on Digital Fundus Images", *Diabetic Medicine*, Vol. 19, pp1105-112.
- [4] Frame, A. J. Undrill, P. E., Cree, M. J., Olson, J. A., McHardy, K. C., Sharp, P. F., Forrester, J. V. (1998). A comparison of computer based classification methods applied to the detection of microaneurysms in ophthalmic fluorescein angiograms, *Computers in Biology and Medicine*, Vol. 28(3), pp225-238.
- [5] Lv, G. Y., Jiang, D. M., Zhao, R. C. (2007). Single Stream DBN Model Based Triphone for continuous speech recognition, *Proceedings of the 9th IEEE International Symposium on Multimedia Workshop*.
- [6] Morizane, K., Nakamura, K., Toda, T., Saruwatari, H., Shikano, K. (2009). Emphasized Speech Synthesis Based on Hidden Markov Models, *Oriental COCOSDA International Conference on Speech Database and Assessments*.
- [7] Yantorno, R. E., Smolenski, B. Y., Iyer, A. N., Shah, J. K. (2004). Usable Speech Detection Using A Context Dependent Gaussian Mixture Model Classifier. *Proceedings of the 2004 International Symposium on Circuits and Systems*.
- [8] Kyoung, K. W., Prugel-Bennet, A., Krogh, A. (2006). Evolving the Structure of Hidden Markov Models, *IEEE Transaction on Evolutionary Computation*, Vol. 10(1).
- [9] Parui, S. K., Guin, K., Bhattacharyam, U., Chaudhuri, B. B. (2008). Online Handwritten Bangla Character Recognition Using HMM, *IEEE Transaction* 2008.
- [10] Baum, L. E., Petrie, T., Soules, G., Weiss, N. (1970). A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains, *Ann. Math. Statist.*, vol. 41, no. 1, pp. 164-171.
- [11] Kwong, S., Chan, C.W., Man, K.F., Tang, K.S. (2001). Optimization of HMM topology and its model parameters by genetic algorithms, *Pattern Recognition*, 34:509- 522.
- [12] Bhuriyakorn, P., Punyabukkana, P., Suchato, A. (2008). A Genetic Algorithm-aided Hidden Markov Model Topology Estimation for Phoneme Recognition of Thai Continuous Speech, *Ninth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing*, 2008.

- [13] Goh, J., Tang, H.L., Al turk, L., Jin, Y. C., (2011). Combining Particle Swarm Optimisation with Genetic Algorithm for Contextual Analysis of Medical Images, 4rd International Conference on Health Informatics, Rome, Italy.
- [14] Ong, Y.S., Lim, M.H., Chen, X. S. (2010). Research Frontier: Memetic Computation – Past, Present & Future, IEEE Computational Intelligence Magazine, In Press, 2010.
- [15] Kennedy, J. & Eberhart, R. C. Particle swarm optimization. Proc. IEEE International Conference on neural networks, Vol. IV, pp. 1942-1948.
- [16] Bakis, R. (1976). Continuous speech word recognition via centisecond acoustic states, Proceedings ASA Meeting, Washington, DC.