A Health Insurance Portability and Accountability Act–Compliant
Ocular Telehealth Network for the Remote Diagnosis and Management
of Diabetic Retinopathy

Yaqin Li, Ph.D.,1 Thomas P. Karnowski, M.S.,2
Kenneth W. Tobin, Ph.D.,2 Luca Giancardo, Ph.D.,2
Scott Morris, M.D.,3 Sylvia E. Sparrow, M.D.,3
Seema Garg, M.D., Ph.D.,4 Karen Fox, Ph.D.,5
and Edward Chaum, M.D., Ph.D.1

1Hamilton Eye Institute, University of Tennessee Health Science
Center, Memphis, Tennessee.
2Oak Ridge National Laboratory, Oak Ridge, Tennessee.
3The Church Health Center, Memphis, Tennessee.
4University of North Carolina, Chapel Hill, North Carolina.
5Delta Health Alliance, Stoneville, Mississippi.

Abstract
In this article, we present the design and implementation of a regional
ocular telehealth network for remote assessment and management of
diabetic retinopathy (DR), including the design requirements, network
topology, protocol design, system work flow, graphics user interfaces,
and performance evaluation. The Telemedical Retinal Image Analysis
and Diagnosis Network is a computer-aided, image analysis tele-
health paradigm for the diagnosis of DR and other retinal diseases
using fundus images acquired from primary care end users delivering
care to underserved patient populations in the mid-South and
southeastern United States.

Key words: diabetic retinopathy, computer-aided diagnosis, tele-
medicine, HIPAA compliance, healthcare outcomes, ocular telehealth,
image analysis, teleophthalmology

Introduction
Diabetic retinopathy (DR) is the leading cause of new-onset
blindness in working-age adults in the industrialized world
today, but timely laser treatments can preserve vision in
patients with DR if the disease is diagnosed at an early stage.
There are currently more than 25.8 million people with Type 1 and 2
diabetes in the United States, with over 7 million being undiagnosed.1
The number of people over the age of 20 with prediabetes is estimated
to be more than 79 million, with an incidence approaching 27% in
those over the age of 65.1 The Centers for Disease Control and Pre-
vention recently estimated that the incidence of diabetes in the United
States will approach 30% by 2050, necessitating the annual screening
of more than 110 million patients for diabetic eye disease. The World
Diabetes Foundation estimates that the number of people with diabetes
worldwide is expected to exceed 438 million by 2030, with 70% of the
cases occurring in low- and middle-income countries with poorly
developed healthcare infrastructure and delivery.2 As an emerging
technology, telemedicine has the potential to be an efficient and cost-
effective way to deliver healthcare services over a large geographic
region and to underserved populations.3–6 With the widespread access
to the Internet and the increasing availability and decreasing cost of
network bandwidth, integration of Internet connectivity into medical
information exchanges can provide high-throughput image analysis
without incurring significant infrastructure costs.

Our approach to ocular telehealth (OT), the remote detection and
management of DR, applies content-based image retrieval (CBIR)
methods to diagnosis of DR based on features present in fundus
images of the retina obtained using commercially available non-
mydriatic cameras. The ultimate goal of our work is to develop and
validate methods to permit more fully automated detection, stratifi-
cation, diagnosis, and management of diabetic eye disease in real
time from digital images taken in a primary care setting. In our OT
network, retinal images from diabetic patients are transmitted in
Digital Imaging and Communications in Medicine (DICOM) format
from fundus cameras to our diagnostic server. Features are extracted
from the retinal images to stratify the level of DR disease using CBIR,
a management plan is recommended, and a report is generated, which
is accessible to the authorized users, including the referring physicians,
the consulting ophthalmologist, and other end user clients.

In this article, the design requirements, network topology, protocol
design, system work flow, and interfaces are described, along with a
brief overview of the image processing and pattern recognition
components of the CBIR diagnostic method. This Web-based net-
work, Telemedical Retinal Image Analysis and Diagnosis (TRIAD),
provides portability, flexibility, and improved efficiency in health-
care delivery by providing access to expert diagnosis of DR in the
primary care space, to meet the growing need for eye disease
assessment and management in rapidly expanding at-risk and un-
derserved patient populations.

Methods
THE TRIAD NETWORK

To achieve the goal of managing diabetic eye disease in OT networks,
an underlying Health Insurance Portability and Accountability Act
(HIPAA)-compliant communication network infrastructure was estab-
lished. In the TRIAD Network, the participating end user clinics are the

DOI: 10.1089/tmj.2011.0004  © MARY ANN LIEBERT, INC. • VOL. 17 NO. 8 • OCTOBER 2011  TELEMEDICINE and e-HEALTH 627
imaging sites. A nonmydriatic fundus camera (not requiring dilation of the pupil) in the clinic captures retinal images from patients during visits to their primary care provider and communicates with the diagnostic server using standard ethernet connections and secure Web-based interfaces. The diagnostic server is the central hub of the communication and data analysis network. The network is also currently designed to permit the integration of semiautomated functionality (e.g., image quality assurance [QA] metrics), image analysis, and management algorithms to assist the consulting ophthalmologist using CBIR methods.

CONTENT-BASED IMAGE RETRIEVAL

In Figure 1, we show the operations our system performs for retinal analysis (under the “diagnostic server” block). Acquired images are inspected for adequate quality using a quality evaluation algorithm and metric.7,8 Images are obtained and stored in a database for consulting physician access using desktop and smart platforms. The anatomical regions of the eye (optic disc, macula, and vessels) are identified.9 Algorithms to detect microaneurysms and exudates are next performed and numerical descriptors of the population of these anomalies are computed.10 The numerical descriptors are rapidly compared with the database of several thousand other image examples.11–14 Matches receive a confidence measure based on the number of examples of different disease in the database to determine the probability of a specific level of retinal disease. Actual diagnostic reports are generated and signed by consulting retinal physicians and returned to the end user clinic in real time.

TRIAD NETWORK ARCHITECTURE

Overview. The TRIAD Network is composed of three components: participating primary care end user clinics (referring physicians), consulting ophthalmologists, and the dedicated diagnostic server. The participating clinics and consulting ophthalmologists are geographically distributed. Figure 2 provides an overview of the geographic distributions of the current participating clinics in the states of Tennessee, Mississippi, and North Carolina. The consulting ophthalmologists are located at Memphis, Tennessee, and Chapel Hill, North Carolina. The diagnostic server is located in the server farm of the University of Tennessee Health Science Center in Memphis, Tennessee, with backup data storage at protected university facilities at a remote site. Network access to deidentified clinical data is also provided to our collaborators at Oak Ridge National Laboratory (ORNL) for the purpose of algorithm development and testing.

Design requirements. To provide a network infrastructure for HIPAA-compliant data collection and communications, the security, reliability, and integrity of the data exchange are crucial in the design specifications. To ensure security, the data transmission protocols must be cryptographic and compliant with Federal Information Processing Standard (FIPS) 140–2, and the patient-sensitive information must be encrypted to meet the HIPAA-compliance requirements.15 To guarantee data reliability, robust data flow in the network is required under adverse network connectivity conditions, and the data transmission must be fast enough to provide high-throughput image analysis and data-processing capabilities. To assure data integrity, validation for data entry needs to be enforced, and the data storage in the system is continuously monitored and regularly backed up to a second secure, remote site, with any changes recorded in audit trails.

Network topology. A hub-and-spoke architecture was selected for the network topology. The hub-and-spoke architecture best serves the needs of data acquisition and management for its flexibility, scalability, and low cost. The hub provides centralized communication and data-processing functionality, thus enabling spoke applications to be removed and replaced without disrupting other spoke nodes. The model offers lower total cost for network establishment and easy scalability when a new spoke site joins the network. The diagnostic server is designated as the hub site, and the end user clinics act as the
spoke nodes, as shown in Figure 1. The server is the logical center for data communications, processing, and management. Inside the server, a dedicated clinical portal and a physician portal are set up for bidirectional data transmission. The dedicated clinical portal provides a channel for each clinic to submit retinal images and metadata captured by a commercial nonmydriatic fundus camera before they are archived in a structured database inside the server for further processing and retrieval. The physician portal not only provides a channel for the consulting ophthalmologist to oversee the generated report for confirmation and validation, but also enables Web-based accessibility to database query for referring clinicians and other authorized users.

In addition to being a communications hub, the server also provides the functional image processing, data archival, and Bayesian decision making for disease stratification and CBIR analysis. Fundus images are numerically evaluated in real time by an automated QA metric at the server after submission and before they are stored in a relational database. The QA assures adequacy of image quality for the purpose of determining a diagnostic assignment. The diagnostic engine then retrieves the images and associated patient metadata from the database, processes the images to identify anatomical structure, extract features, and detect lesions. Following the analysis, the server assigns a tentative diagnosis with an appropriate management plan for the stratified level of disease and generates an encrypted E-mail alert for the consulting ophthalmologist for confirmation. The ophthalmologist reviews the images, issues a final report with diagnosis, and provides a validated report to the end user using a signature certificate, which is returned via encrypted PDF and password-protected E-mail attachment or directly deposited into the electronic health record.

Protocols design. The data exchanges in the OT network are structured in various formats; for example, the retinal images submitted from a clinical spoke site are in DICOM file format, along with defined metadata in text file. The committed storage of the data entries inside the server are formatted into relational database records, the diagnostic report generated on the server is PDF file, and the physician portal presents the data for authorized access in health level 7 standards. For data security in the network, all data transmission protocols are cryptographically designed to meet the FIPS140-2 standards for HIPAA compliance. An overall protocol design is illustrated in Figure 3.

Secure file transfer protocols (SFTPs) are used for all file transfers to provide authorized data submission and encrypted file transfer between the clinic site and the server. Specifically, SFTP is used between the end user and clinical portal to transfer image files to the server and communicate the QA result (pass or fail) to the end user. After the completion of the image processing, data analysis, and validation, a report is generated in a password-encrypted PDF document for secure access through a Hypertext Transfer Protocol over Secure Socket Layer (HTTPS)-secure Web site provided by the physician portal or in encrypted E-mail attachment. To access the report generated on the server, the authorized users can interact with the physician portal in two ways: either login on a secure Web site using HTTPS or download an encrypted PDF report sent as an E-mail attachment. To ensure data security and integrity during report generation, the ophthalmologist digitally signs and encrypts the report using an X.509 certificate at the end of the reviewing and confirmation process.

To ensure data security during archival and retrieval, all data access is controlled and validated, and all data archival processes are encrypted. To submit the clinical data and the images into the server, the SFTP issues password-controlled access to the server and a dedicated secure key validation. To archive the data entries into the database or to retrieve data records, password-controlled access is required through a secure connection to the database. To ensure data security under adverse network conditions or unstable network connectivity, the spoke application persists on a robust data transmission with no-data-loss guarantee. A persistent acknowledgement is required for committed storage in the bidirectional data transmission between the hub and the spokes to assure there is no loss of data. The client program on the camera closely monitors the network status in a real-time manner and reacts.

Fig. 2. Geographic distribution of the participating clinics in the TRIAD Network. The TRIAD Telehealth Network manages diabetic retinopathy assessment and management for working poor urban patients in Memphis and Shelby County, TN (Plough Foundation Network), and underserved communities in rural Mississippi through the Delta Health Alliance, TEAM (Training, Education, Access, and Management) Sugar-Free project in the Delta and North Mississippi. We also partner with primary care clinics affiliated with the University of North Carolina. Colleagues at Oak Ridge National Laboratory develop, test, and validate new diagnostic algorithms using deidentified retinal images from the server.
promptly in case of connectivity instability. In the multithread client monitoring program, one of the threads is devoted to scanning the network status constantly and reporting to the end user at the spoke site on a continuous basis. If connectivity issues arise, the spoke application attempts connection recovery if possible and alerts the end user otherwise. If data transmission is in process while network problems occur, the client program reattempts transmission of the exported images after the network is reinstated.

**Work flow.** After the architectural design of the network topology and protocols was put in place, the next step was a detailed procedural design of the system workflow. As illustrated in Figure 3, the functional modules are categorized into three different groups: data storage modules, procedural modules, and graphical user interfaces (GUIs).

Data storage can be either in file format inside directories (shown as folders) or in a database format (shown as cylinders). The images are stored as DICOM and Joint Photographic Experts Group (JPEG) files inside the folders designated for each clinic. There are three major folders corresponding to different stages of image processing: the incoming folder for the raw images, the prevalidation directory for the intermediately processed result, and the expert-validated directory for validated items mature enough (or ground truthed) to be incorporated into the CBIR library. There are two interrelated databases in the system, one for structured organization of the patient information and the examination of clinical data for future diagnostic decision making and query, and the other for CBIR indexing to achieve automated diagnosis. The images are interconnected to both patient archival and CBIR databases, respectively, by link associations with an individual exam semantically or to a numerical CBIR record. As more quality images are acquired into the system, the two databases grow in parallel, and the more enriched data corpus can provide better statistic precision in the Bayesian decision making for the purpose of more fully automating the diagnostic process.

**Graphical user interface design.** GUI modules provide the end user with interactive access to the telemedicine network and enable monitoring of the information exchanges between the data storage and the procedural modules. The GUI design in the network comprises three parts: (1) the application GUI at spoke site for the end user clinics, (2) the Web-based physician GUI for the authorized referring physicians and the consulting ophthalmologist, and (3) the service GUI for monitoring and bookkeeping. A transparent service GUI automatically runs on the server for system monitoring by the administrator.

The client application GUI provides an interface for the end user to submit the data entries and monitor the network status. The clinic acquires fundus images of the retina using a nonmydriatic camera (VisuCam Pro NM; Carl Zeiss Meditec, Thornwood, NY). Newly exported images are automatically encrypted and transmitted to the hub site through the client GUI along with metadata entries from the clinics. At the same time, the client application GUI continuously scans the network status and reacts in a real-time manner to prompt interactive information if and when connectivity issues occur.

The server Web GUI provides an interface for authorized access of the diagnostic result. For the consulting ophthalmologist, the images and patient metadata are reviewed through the Web GUI for digital signature, encryption, and report to the referring physician. For the referring physician, the Web GUI provides a secure access to the diagnostic result along with patient information. The server Web GUI can be accessed through a fixed or mobile smart device, provided that it has browser access to the Web. A Web GUI for recommendation and confirmation from a consulting ophthalmologist is shown in Figure 4.
lesions in the fundus image, and disease stratification based on CBIR constructed from the numerical features.

Quality assessment. The QA method created for our network is computationally efficient and thus enables the network to provide high-speed quality check to the end user. In the method, a segmentation of the vascular tree is performed and analyzed by breaking the vessel pixels into wedge-shaped and annular regions, the color content of the image is measured and the intermediate results are then combined to build a complete feature vector that describes the quality-essential image aspects. Comparison with a training library using machine-learning methods results in a quality measurement, which has been shown to correlate well with a physician’s ground-truth assessment of image quality for the purpose of diagnosing retinal disease.7,8

Anatomical features of the retina. Anatomical features of the retina (vascular tree, optic disc, and macula) are required to evaluate the image quality and locate key structures used in diagnosing disease in the retina. The main purpose of segmenting the vessel structure is to evaluate quality and to serve as landmarks for locating the optic disc and macula.7 Our implementation is fast with an average processing time of < 1 s using an off-the-shelf PC. The optic disc and macula detection uses our published algorithm.9 In this method, four features of the retina (brightness, vessel thickness, density, and orientation) form a Gaussian-based probability density function, which models optic disc regions and non-optic nerve regions. The estimated position of the optic disc then serves as a parameter for a parabolic fit to the vascular tree, which is estimated using nonlinear least-squares algorithm to fit the orientation and curvature of the vessel structure. The estimated orientation and curvature are used to “point” to the macula position, which is then obtained by using the mean value of the distance between the optic disc and macula derived from a training set.

Lesion segmentation. Our system focuses on the detection of two major lesions associated with DR: microaneurysms and exudates, although our methodology is extensible to other retina diseases such as age-related macular degeneration, provided the CBIR library is populated with sufficient examples. Microaneurysm detection utilizes a wavelet-based background removal process and a “Radon Cliff” operator, which detects Gaussian-like circular structures regardless of their size or position in a window.16 Exudates are detected using a median-filter based background-subtraction method followed by dual measurements performed using Kirsch’s edges and Haar wavelet.17 Using the exudates methods in a publicly available dataset containing normal images and images with clinically significant macular edema resulted in a sensitivity of 100% and specificity of 50%, which is a potentially powerful screening tool for diabetic macular edema.17

Disease diagnosis using CBIR. Detected lesions are grouped together for each image and a complete description of the image is created based on the lesion population and additional global texture measurements, which are measured on the entire fundus image. This “lesion population vector” constitutes a query, which is compared with a database of past images with recorded diagnoses. The comparison returns a number of “nearest neighbors” using an efficient database search, which constitutes an estimate of a posterior probability, similar to a classical kNN classifier, but in our case a weighted sum of the similarity between query and returned images are computed to obtain better sensitivity.12 A confidence value based on Poisson statistics is also computed, which creates a mechanism for estimating the accuracy of the retrieval and possibly invoking “physician oversight” in an automated system. Our published results include experiments on 1,355 macula-centered images obtained from a DR screening program in The Netherlands1,4 and a smaller set of 98 images from a Native American population.18 Using our methods, a sensitivity and accuracy of 90% and 95%, respectively, were obtained.

Image processing. The image-processing procedural modules are run in an automated fashion inside the server, addressing all the data...
processing and managements aspects of the network. Most of the procedural modules are implemented as constantly running services transparent to the end users. The images together with the associated patient metadata first go through a QA module to check data consistency, assure correct anatomical location, and issue image QA metric. The QA method implemented in our network provides a high-speed quality check to the clinical operator. An estimate of the quality on a scale from 0 to 1.0 is made, and if the numeric value is below a threshold, the image is rejected as poor quality. By thresholding the numerical QA metric, the end user gets a token of “Pass” or “Fail” for the quality check. The quality metric has been shown to correlate well with a physician’s ground-truth values.7

The incoming data that pass QA are then forwarded to the diagnostic engine to perform feature extraction, lesion detection, and CBIR-based disease stratification according to the posterior probability of each defined disease state. Bookkeeping functionality is another integrated part of the procedural modules. Bookkeeping provides a window for monitoring the data flow and plays an important role in maintaining system reliability and data integrity.

Reliability and data integrity. Reliability and data integrity are critical in the procedural design of the system workflow. Reliability requires the completeness, timeliness, and accuracy of the data, whereas integrity refers to the validity and consistency of the data. In a real-world system, reliability and data integrity can be compromised in a number of ways, for example, human errors when data are entered manually, errors that occur during data transmission, virus attacks, hardware malfunctions, and even natural disasters.

To minimize these threats, several precautionary methods are incorporated in the system design: regular backup to an independent physical media in the university server farm, control of data access via the described security mechanisms, user interfaces designed to prevent the input of invalid data, and bookkeeping of the critical status of every procedure or GUI involving data manipulation to provide an audit trail of data flow within the system.

Regular backup provides data recovery in case of system failure and therefore is important to enable data availability for continued use. In our network, the backup routine is scheduled on a daily basis and performed at two different levels: backup of the data and backup of the system. The data to be backed up include the database and the relevant directories in the hard drive. The data drive of the server is set up on an incremental backup daily, capturing the new data that have been created and tracking changes to the data. In addition, the entire server system is backed up every day, permitting a total restoration from scratch if necessary. Two copies of the backups are stored in independent media and physically kept at two geographically separated locations.

Data access control is another security measure to guarantee system safety while granting authorized access to the system. Each channel to interact with the network system is controlled by predefined username and password combination. Also, every connection to the hub is encrypted and requires validation; for example, the SFTP access from the clinics to the designated folders on the server, the HTTPS access from the physicians to the recommendation and management plan, and access to the Structural Query Language (SQL) database on the server are only granted to authorized users. Validation with issued password or dedicated certificates is required for any data exchange between the spokes and the hub.

To prevent invalid data entry from the end user, the interactive GUI is designed to provide syntactic data validation before submission, and data consistency check is performed before they are deposited into the database or the directories on the server. To keep track of the changes in the system, including the usage, data access, abnormalities, etc., bookkeeping provides a meticulous audit trail of the whole system in various ways: a log is kept on every procedure on the server, either dedicated in the format of event log for each procedure, or in text document recording the changes of every critical state in the procedures. Any connection to the database and changes...
applied to the records, whether initiated by the end user or by the autonomous process on the server, is also recorded in an audit table in database format.

PERFORMANCE EVALUATION

Geographic and demographic domain performance. The TRIAD Network currently has three regions of active operation: in urban Memphis, Tennessee, in rural counties of northwest Mississippi (Delta Health Alliance), and in Chapel Hill, North Carolina. Since its initial deployment in February 2009 at the Church Health Center (a health center for the working poor in Memphis, Tennessee), the network has provided diagnostic reports on 5,586 retinal images collected from 2,892 exam sessions on 2,510 patients. Categorized in terms of geographic locations, 854 of the exam sessions, comprising 29.5% of the total, are from 676 patients in Memphis (including some 1-year follow-up images), 178 of the exam sessions (6.2%) are from 144 patients served by Delta Health Alliance in rural counties in northwest Mississippi and 1,860 exam sessions (64.3%) are from 1,691 patients in the region surrounding Chapel Hill, North Carolina.

Categorized by disease severity and stratification, 1,883 (75.0%) patients had no evidence of DR or other retinal diseases. These patients will be managed in the primary care clinic with follow-up retinal photography in 12 months. The incidence of any DR in the population was 14.8%. Of these patients with any DR, 64.5% of eyes (9.6% of the patient total, 362 patients) had minimal or mild DR without evidence of fluid leakage (clinically significant macular edema) and were not retested for a formal eye exam. These patients will be reevaluated in the primary care clinic in 6–12 months (12 months for patients with only rare [≤5] microaneurysms). In summary, 82.6% of all diabetic patients examined to date will continue to be managed in the primary care setting. A total of 294 (5.3%) of all 5,586 retinal images evaluated had DR severe enough to warrant referral and treatment. An additional 370 retinal images from 252 patients (10.0%) had retinal findings exclusive of DR that warranted a formal ophthalmic evaluation. These findings included evidence of macular degeneration, vascular occlusions, and optic disc findings suggestive of glaucoma, among others.

Time domain performance. Rapid data transmission and high throughput is one of the key features of the TRIAD Network. To validate the performance of the designed network, the average time from image acquisition to report generation is applied as a metric for time domain performance evaluation. The system response time was measured by observing 60 sessions of data transmission between a fundus camera in an active clinic and the diagnostic server. As findings may vary in different images with different disease state, the time response for different categories of images may vary. The average response time from image export to the generation of a diagnostic and management report for the end user is approximately 113.8±19.9 s when validation is performed by the consulting ophthalmologist in real time (Figure 5). The network structure provides sufficiently high throughput for real-time image analysis and diagnosis of diabetic eye disease.

Summary

There is significant interest and ongoing research into the application of Web-based telemedical networks for the diagnosis and management of DR. More recently, these networks have been employed to test and validate more fully automated methods for screening large, at-risk populations. Our TRIAD Network expands upon these capabilities by integrating real-time QA metrics into retinal image acquisition and creating an infrastructure to manage reporting in real time from mobile platforms and fixed medical care sites in an HIPAA-compliant manner.

The TRIAD Network provides a method for low-cost, real-time diagnosis and patient referral in the primary care environment, providing access to expert diagnosis for underserved patients, and high-throughput methods to meet the growing need for screening in rapidly expanding at-risk populations. The responsive speed of the network is demonstrated to achieve near real-time capability for image processing and data analysis. The design of the network using existing Web-based infrastructure and commercially available nonmydriatic cameras provides easy scalability for broad geographic implementation.

Acknowledgments

These studies were supported in part by grants from ORNL, the National Eye Institute (EY017065), and the Health Resources and Services Administration through a grant to the Delta Health Alliance, by an unrestricted UTHSC Departmental grant from Research to Prevent Blindness, New York, NY, and by The Plough Foundation, Memphis, TN. Dr. Chaum is an RPB Senior Scientist.

Disclosure Statement

Drs. Chaum and Tobin are principals in a startup company based upon the content-based image retrieval (CIBR) methods described. The intellectual property and patent application are the property of University of Tennessee-Battelle Corporation.

REFERENCES


Address correspondence to:
Edward Chaum, M.D., Ph.D.
Hamilton Eye Institute
University of Tennessee
930 Madison Ave., Suite 731
Memphis, TN 38163

E-mail: echaum@uthsc.edu

Received: January 6, 2011
Revised: March 8, 2011
Accepted: March 9, 2011
This article has been cited by:

