Observer Detection Performance in Radiology Using a Retransmission-Free Network Communication Protocol

Elizabeth A. Krupinski, Ph.D. & Yasser Alsafadi, M.S.

Department of Radiology, The University of Arizona

Reprint requests should be sent to : Elizabeth A. Krupinski, Ph.D., Department of Radiology, Radiology Research Division, The University of Arizona Health Sciences Center, Tucson, AZ 85724. Phone: 602-626-4498. Fax: 602-626-4376.

ABSTRACT

RATIONALE AND OBJECTIVES To measure the effect of image data loss on diagnostic accuracy in order to test the possibility of using a retransmission-free network communication protocol for transferring radiologic images.

METHODS. Mammograms transferred over a simulated network with 0%, 15% and 25% transmission packet loss were presented randomly to 10 observers who typically read mammograms. Observers reported on the presence or absence of microcalcification clusters, and the number of calcifications per cluster.

RESULTS. Performance with 15% loss did not differ from performance with 0% loss. 25% loss resulted in a significant decrease in performance. Accuracy of counting individual microcalcifications was quite high in all 3 loss conditions.

CONCLUSIONS. Up to 15% packet loss can be tolerated without affecting diagnostic performance. These data suggest that, in some radiologic applications, retransmission-free communication protocols may be useful.

KEY WORDS. Retransmission-free communication protocol; image data loss; observer performance; visual perception.

Observer Detection Performance in Radiology Using a Retransmission-Free Network Communication Protocol

Radiology is at the beginning of a new era - an era in which digital imaging, teleradiology, and PACS (Picture Archiving and Communications Systems) are becoming more and more commonplace. With the implementation of these new digital technologies comes a host of challenges - how to acquire, store, retrieve, transmit and display massive amounts of image and text data.[1-3] From the radiologist's point of view, data retrieval and transmission must be quick and diagnostic accuracy using these digital images must be at least as good as with traditional film images.

There are many ways in which image quality (hence, diagnostic accuracy) can be compromised and image data lost or corrupted during acquisition, storage, retrieval, transmission and display procedures.[4-10] With respect to image transmission, the ACR-NEMA Working Group 6 is working on the DICOM Standard, which deals with what types of networking protocols should be used for transmission of medical images.

Currently, the TCP/IP Internet is the most widely used network. In this network, the low-level protocol (IP) supplies only best-effort delivery - bandwidth is not reserved, and if a gateway is congested packets are discarded without regard for requirements of competing connections. Because of this design, it is possible to lose packets - on the order of 10% or more - if transmission is aggressive (i.e., lots of data sent at high speeds). Additionally, loss is often bursty in nature, i.e., two or more back-to-back packets are lost at a time. Recent measurements of these loss bursts on the TCP/IP Internet indicated an average loss burst of 3.3 packets with 12% of the losses having a burst size of greater than four packets.[11]

To make the data delivery (IP level) more reliable, the high-level TCP protocol uses retransmission to recover lost packets. While retransmission insures lossless data delivery, avoiding retransmission could offer two advantages. The first advantage is an improvement in end-to-end response time. For example, a 1Gbps network can transport a 10Mbyte image in approximately 100ms.[11] If only one packet is lost and needs to be retransmitted, response time doubles on a wide-area network with 100ms round-trip latency. If more packets are lost, then multiple round-trips may be needed to transmit all the image data successfully.

The second possible advantage of a retransmission-free protocol is that it would reduce the buffer space required by the sending machine. Protocols that retransmit lost data packets have to buffer all transmitted data until receipt has been acknowledged. For example, using a 1Gbps network with 100ms round trip latency, TCP would have to either reduce its window size (i.e., reduce its effective throughput rate), or maintain 12.5 Mbytes of buffer space per established connection.[11]

Given that speed is important for the transmission of radiologic images in many circumstances (e.g., teleradiology, consultation from a home-based PC via modem, battlefield conditions), the purpose of the investigation presented here was to explore the possibility of using a retransmission-free or light-weight[12] communication protocol (i.e., a protocol designed to minimize the overhead for data transmission and reception) for transferring radiologic images. Of primary interest was the perceptual/performance issue - can diagnostic accuracy be maintained if some amount of image data is lost during transmission.

METHODS & MATERIALS

Encoding and Recovery Schemes

In this experiment, we used the scattered-encoding mechanism described by Turner and Peterson.[11,13] The technique is designed for use with retransmission-free communication protocols, and aims to minimize the impact of packet loss by encoding the images in such a way as to insure that the pixels contained in a given packet are scattered throughout the image. Thus, if a packet is lost redundancy in the remaining image data can be used to reconstruct the lost information. This technique has also been well characterized with respect to its influence on compression, what types of reconstruction algorithms are most efficient, computation time and end-to-end performance.[13]

The encoding scheme used in the present experiment is derived directly from Turner and Peterson.[11] Given an image of size RowLength x ColumnLength pixels, regular-scattered encoding (assuming a byte per pixel) is expressed using two key parameters: ByteOffset and PacketOffset. ByteOffset controls how bytes within a single packet are distributed throughout the image. It specifies the number of bytes between adjacent pixels in the same packet (ByteOffset = RowLength + 2). It is this parameter that ensures that no two bytes within the same packet are ever adjacent in image space.

PacketOffset addresses the problem of bursty packet loss. It specifies the number of bytes between starting pixels of adjacent packets (PacketOffset = 16). This parameter controls how far apart temporally adjacent packets are in image space. The result of this type of encoding scheme is that packet loss appears on the image as a series of non-adjacent diagonal lines. In the present experiment, packet size was chosen to be the expected network maximum transfer unit (MTU), i.e., 1024 Bytes. The encoding and network packets were the same.

There are a range of available recovery options (used once the transmitted image is received) which are ordered by increasing accuracy and processing requirements.[14,15] We chose to use a four-neighbor interpolation recovery operation (i.e., once the image is transferred, intensity values are assigned to lost pixels by averaging the four nearest neighbors and assigning that value to the lost pixel).

Study Images

In order to test the effects of packet loss on observer performance, a set of images was needed containing a lesion likely to be distorted or even eliminated from the image if loss occurs. Therefore, we chose to use mammographic images with subtle microcalcification clusters. Clinically significant microcalcifications are considered to range in size up to 0.75 mm in diameter with most less than 0.5 mm,[16,17] and thus could be eliminated with packet loss.

A test set of 50 mammographic images was collected from the Tucson Breast Center and the Mammography Department at The University of Arizona Health Sciences Center. Twenty-five lesion-free images and 25 images containing a single, subtle microcalcification cluster were selected. Only one image per case (i.e., right or left breast, craniocaudal or mediolateral projection) was used. The test set represented a typical range of dense to very dense breast parenchymal patterns. Since the images were obtained from past records, cluster locations and all diagnoses for the cluster-containing images were confirmed by a surgical pathology report. The lesion-free images were obtained from past records and an existing film library, both of which contained patient records verifying the lesion-free status of the patient for at least three years.

The 50 test images were digitized using a Lumisys film digitizer (Lumisys Inc., Sunnyvale, CA) with a spot size of 50 μm and a contrast resolution of 12 bits. Images were transferred with no loss via Ethernet network to a DEC VAX 8600 Computer (Digital Equipment Corp., Maynard, MA) for processing. Burst-oriented packet losses were then incurred for each image by encoding the images using the protocol described above and then uniformly dropping packet groups over the entire image area. Group size was normally distributed with a mean of 8 and standard deviation of 1 packet. The four-neighbor linear interpolation recovery scheme was then run on each image.

The amount of loss to simulate was chosen to be a "worst-case scenario". This was done 1) so that the amount of packet loss was very visible on the image (i.e., maximally likely to influence detection performance), and 2) to show that if observer performance is not influenced significantly at extremely (even unlikely) high loss rates, then performance should not be influenced at much lower, realistic loss rates. Therefore, 0%, 15% and 25% packet losses were used, resulting in a final test set of 150 images. Figure 1 shows an example of a breast image with 25% loss before and after recovery. Visual examination of each film (with loss but before recovery) confirmed that at least three packets of image data were lost within each microcalcification cluster.

Insert Figure 1 Here

The 150 images were then written to film using a Kodak Ektascan Laser Printer (KELP model 100, Eastman Kodak, Rochester, NY), which supports 12 bits of grey level and can print a matrix size of 4084 x 4984 pixels on 14" x 17" film. The images were written onto DuPont helium neon laser imaging film (4 images per film, and there was no magnification or distortion of the images), and developed by a Kodak M35A X-OMAT processor (Eastman Kodak, Rochester, NY). Mean optical density and contrast of the processed images were carefully controlled to match closely those of the originals.[18]

Procedure

An experienced staff mammographer not participating as an observer viewed the 50 0% loss images. The task was two-fold. The first part was to search each image for the presence or absence of microcalcification clusters and localize them on an outline of the breast provided for this purpose. This served to corroborate the lesion-free or lesion-containing status of each image as designated in the official report and to confirm the cluster locations. The second task was to count the number of individual microcalcifications per cluster visible on each image. There were four categories for number of microcalcifications : 1-5, 6-10, 11-15 and > 15. Five clusters contained 1-5 microcalcifications, 7 contained 6-10, 7 contained 11-15, and 6 contained > 15 microcalcifications. The 0% loss image was used as the baseline condition. The original film image was not used to count microcalcifications because the processes of digitizing and writing to film already incurred some degree of loss. By using the 0% loss film for the baseline condition, all other sources of loss were equalized except for the simulated packet loss.

The 150 test images were randomized and divided into three blocks of 50 images each, with the restriction that only one version (0%, 15% or 25% loss) of a given image could appear in a block. A different randomized order was used for each reader. Six experienced staff radiologists who typically read mammograms and four radiologic technicians with advanced certification in mammography served as observers. In three half-hour sessions separated by four weeks on average, the readers viewed the three blocks of test images. A period of four weeks should have been sufficient to promote forgetting of images[19] since this sort of time frame is used in many ROC studies using a repeated measures design,[20-22] especially given the fact that the observers in the present study saw hundreds of different mammographic images within that time.

The task was to search each image thoroughly for the presence of microcalcification clusters and report a decision in 3 parts: 1) cluster present or absent, 2) confidence in that decision using a 5-level rating scale where 5 = cluster present, definite and 1 = cluster absent, definite, and 3) how many microcalcifications are present in the cluster. The observers used the same microcalcification counting scheme described above for the baseline determination procedure.

RESULTS

The decision confidence ratings were used to generate individual receiver operating characteristic (ROC) curves and compute area under the curve (Az) and standard deviation values for each reader in the three test conditions (0%, 15% and 25% loss). The ROC results and Analysis of Variance (ANOVA) statistical analysis were carried out using the multireader-multicase (MRMC) method described by Dorfman, Berbaum and Metz.[23]

A preliminary analysis which compared the performance of the six radiologists with the performance of the four radiologic technicians revealed that there were no statistically significant differences between the two groups of observers. Therefore, for all subsequent analyses the data from the ten observers were treated as a single group.

The individual Az values for the ten observers are given in Table 1. The ANOVA analysis revealed a significant main effect for treatments (i.e., % loss) (F = 2.12, $\underline{p} < .001$) and for cases (i.e., images) (F = 2.41, $\underline{p} < .01$); and a significant interaction effect for cases x readers (F = 3.86, $\underline{p} < .001$). Post-hoc tests (Fisher Protected Least Squares Difference) indicated that the average difference in Az (0.03 in favor of 15% loss) between the 0% and 15% loss conditions was not statistically significant, while Az in the 25% loss condition was lower and differed significantly from both the 0% (Az difference = .04) and 15% loss (Az difference = .07) conditions ($\underline{p} < .05$).

Insert Table 1 Here

Table 2 presents the percentage of true-positive and false-positive reports for each observer in the three loss conditions. The decrease in Az performance in the 25% loss condition was due primarily to an increase in the percent of false-positive reports, although a slight decrease in the percent of true-positive reports was also noted. A repeated measures ANOVA indicated that there was a significant main effect due to treatment for the percentage of true-positive reports (_F_ (2,18) = 6.27, $p < .009$). Post-hoc tests indicated that the percentage of true-positive reports for 0% loss did not differ significantly from either 15% or 25% loss; but that 15% loss did differ significantly from 25% loss ($p < .05$). The pattern of results was the same for false-positive reports (_F_ (2,18) = 4.46, $p < .03$).

Insert Table 2 Here

The final analysis dealt with the number of individual microcalcifications counted in each cluster. For this analysis the four categories for the number of microcalcifications were used : 1-5, 6-10, 11-15, and > 15. The correlation between baseline judgements for each image and the judgments made by the observers was then determined for each of the ten observers in the three loss conditions (0%, 15%, 25%). False-negatives were not included in the correlation analysis. Including the false-negatives (i.e., using zero for the counted number) would have decreased the correlation coefficients slightly across all three conditions for each of the observers. The coefficients of correlation for the ten observers are presented in Table 3.

Insert Table 3 Here

It can be seen from Table 3 that overall the correlation between the actual number of microcalcifications per cluster and the judged number per cluster was high (range = 0.73 - 0.97) for all subjects in all three loss conditions. A subsequent examination of the microcalcification counting data indicated that when there was disagreement between the baseline and observer judgments, the observer was off by only one category (e.g., baseline = category 1, judgment = category 2) in 83% of the cases and by two categories in the other 17%. In 67% of the cases the discrepancy was between categories 3 (11-15) and 4 (> 15).

DISCUSSION

The results of this experiment satisfactorily addressed our primary interest in conducting the study - from a perceptual point-of-view, diagnostic accuracy and the ability to identify fine image details (i.e., individual microcalcifications) can be maintained if some amount of image data is lost during network transfer.

Aside from the perceptual question of whether packet loss influences observer performance, is the question of whether it would be practical to actually implement and use a light-weight retransmission-free protocol in radiology. As noted previously, two advantages of a retransmission-free protocol are that 1) response-time is greatly improved, and 2) buffer space can be reduced on the sending machine.

A number of other issues are relevant to the question of practicality as well. One question is whether or not images can still be compressed using the encoding scheme described here. The answer is yes. Turner and Peterson[13] tested the scattered encoding algorithm to determine what impact it had on three widely used types of compression algorithms : run-length encoding (RLE), differential pulse code modulation (DPCM) and discrete cosine transform (DCT). They found that it was possible to compress images (after encoding) using all three methods and that compression ratios were on average about 80% as efficient as those attainable using the full unencoded images. Use of the family of compression schemes based on the widely-used Lempel-Ziv algorithm can also be used with this encoding scheme.[24]

Another issue is computation time and end-to-end performance. The encoding time is negligibly different from any other encoding scheme.[11] Once encoded images can be stored, and since the encoded packets are the same as the network packets data can be copied directly from the storage disk to the network interface. Computation time for the four-neighbor interpolation recovery is also minimal. Turner and Peterson[13] compared computing times for no-recovery vs. four-neighbor interpolation recovery. No-recovery takes little or no time. At a packet loss rate of 1%, four-neighbor interpolation processes an image at a rate of 27.2 MBytes/sec (36ms for a 1 MByte image); and at 10% loss, it processes an image at 2.7 MBytes/sec (360ms for a 1 MByte image).

Finally, with respect to end-to-end performance, Turner and Petersen[13] have compared this light-weight protocol with FTP (file transfer protocol) over the Internet (1.5Mbs and 200-300ms average latency) and found that the light-weight protocol performs two to three times better than FTP.

In conclusion, the present experiment demonstrated effectively that diagnostic accuracy does not seem to be affected significantly by image data lost during packet transmission. This suggests that if a retransmission-free protocol was adopted, it might be practical to use in select clinical situations : teleradiology to remote areas that must use telephone lines rather than the Internet; the use of PCs and modem hook-ups for use by radiologists consulting after-hours while at home; fax transmission of images[25] and teleradiology in battlefield conditions[26-28]. As noted previously, the amount of data loss chosen in this study was an extreme, worst-case-scenario. Such losses are very unlikely to be encountered in the real world. The results do suggest, however, that if extreme loss (as in this study) does not seem to affect diagnostic accuracy, then lower rates of loss (which may be encountered) should not affect diagnostic accuracy either.

REFERENCES

1. Hindel R. Digital image storage technology. Invest Radiol 1993;28:454-458.

2. Zink S, Jaffe CC. Medical image databases : a National Institutes of Health workshop. Invest Radiol 1993;28:366-372.

3. McGarty TP. Multimedia communications technology in diagnostic imaging. Invest Radiol 1991;26:377-381.

4. Seeley GW, Fisher HD, Stempski MO, Borgstrom M, Bjelland J, Capp MP. Total digital radiology department : spatial resolution requirements. AJR 1987;148:421-426.

5. Seeley GW, de Valk JPJ, Kroon HM, Rompelman O, Bakker AR. Image compression evaluation m: an example of a PACS component analysis chain using psychophysics. SPIE Medical Imaging II : Image Data Management and Display 1988;914:792-798.

6. Stewart BK, Dwyer SJ. Prediction of teleradiology system throughput by discrete event-driven, block-oriented network simulation. Invest Radiol 1993;28:162-168.

7. Collins CA, Lane D, Frank M, Hardy ME, Haynor DR, Smith DV, Stewart BK, Parker JES, Bender GD, Kim Y. Design of a receiver operating characteristic (ROC) study of 10:1 lossy image compression. Proceedings of SPIE Medical Imaging 1994 : Image Perception 1994;2166:149-158.

8. Cook LT, Insana MF, McFadden MA, Hall TJ, Cox GC. Assessment of low-contrast detectability for compressed digital chest images. Proceedings of SPIE Medical Imaging 1994 : Image

Perception 1994;2166:159-169.

9. Kostas TJ, Sullivan BJ, Ansari R, Giger ML, MacMahon H. Adaptation and evaluation of JPEG-based compression for radiographic images. Proceedings of SPIE Medical Imaging 1993 : Image Capture, Formatting, and Display 1993;18997:276-281.

10. Lee H, Kim Y, Rowberg AH, Frank MS, Lee W. Lossy compression of medical images using prediction and classification. Proceedings of SPIE Medical Imaging 1993 : Image Capture, Formatting, and Display 1993;18997:282-287.

11. Turner CJ, Peterson LL. The effects of transfer encoding on image quality. Proc. of the 2nd Intl. Conf. on Image Processing, Sept. 1992; 62-67.

12. Doeringer WA, Dykeman D, Kaiserswerth M, Meister BW, Rudin H, Williamson R. A survey of light-weight transport protocols for high-speed networks. IEEE Transactions on Communications 1990;38:2025-2039.

13. Turner CJ, Peterson LL. Image transfer: an end-to-end design. Proc. of the SIGCOMM '92 Symposium, Baltimore, MD, Aug., 1992.

14. Bates RHT, McDonnell MJ. Image restoration and reconstruction. Oxford, England : Clarendon Press; 1986.

15. Pratt WK. Digital Image Processing. New York : John Wiley and Sons; 1978.

16. Sickles EA. Breast calcifications: mammographic evaluation. Radiology 1986;160:289-293.

17. Ackerman LV, Gose EE. Breast lesion classification by

computer and xeroradiograph. Cancer 1972;30:1025-1035.

18. MacMahon H, Vyborny CJ, Metz CE, Doi K, Sabeti V, Solomon SL. Digital radiography of subtle pulmonary abnormalities : an ROC study of the effect of pixel size on observer performance. Radiology 1986;158:21-26.

19. Metz CE. Some practical issues of experimental design and data analysis in radiological ROC studies. Invest Radiol 1989;24:234-245.

20. Mannino DM, Kennedy RD, Hodous TK. Pneumoconiosis : comparison of digitized and conventional radiographs. Radiology 1993;187:791-796.

21. Hoffmann KR, MacMahon H, Doi K, Metz CE, Yao L, Abe K. Evaluation of an enhanced digital film-duplication system by receiver operating characteristic analysis. Invest Radiol 1993;28:1134-1138.

22. Elam EA, Rehm K, Hillman BJ, Maloney K, Fajardo LL, McNeill K. Efficacy of digital radiology for the detection of pneumothorax : comparison with conventional chest radiography. AJR 1992:158:509-514.

23. Dorfman DD, Berbaum KS, Metz CE. Receiver operating characteristic rating analysis : generalization to the population of readers and patients with the jackknife method. Invest Radiol 1992;27:723-731.

24. Peterson LL. Personal communication, June 1994.

25. Sistrom CL, Gay SB. Facsimile transmission of radiographic images : preliminary experiments with a personel computer and a fax modem board. Invest Radiol 1993;28:860-867.

26. Rayman RB. Telemedicine : military applications. Aviation, Space, and Environmental Medicine 1992;February:135-137.

27. Cawthon MA, Goeringer F, Telepak RJ, Burton BS, Pupa SH, Willis CE, Hansen MF. Preliminary assessment of computed tomography and satellite teleradiology from Operation Desert Storm. Invest Radiol 1991;28:854-857.

28. Dao HNV, Cawthon MA, Simmons GE. The use of teleradiology in military medical field exercises. In JM Boehme, AH Rowberg, NT Wolfman (Eds). SCAR 94 : Computer Applications to Assist Radiology 1994; Carlsbad, CA, Symposia Foundation:286-289.

Acknowledgments

We would like to thank Drs. Laurie Fajardo, Cheryll Hicks, Rebecca Hulett, Rebecca Hunt, Pamela Lund and Tad Tanoura for serving as readers in this study.

Table 1. Az values for 0%, 15% and 25% loss. 95% confidence limits in parentheses.

Observer	0% Loss	15% Loss	25% Loss
1	.9509	1.00	.9162
2	.9000	.9319	.8492
3	.8158	.9028	.7838
4	.8927	.8787	.8533
5	.7858	.8378	.8012
6	.8994	.9170	.8575
7	.9566	.9719	.8927
8	.7677	.8532	.8191
9	.8167	.7954	.6683
10	.8243	.8380	.8007
Mean	.8610 (-.8127,+.9092)	.8941 (-.8461,+.9421)	.8242 (-.7749,+.8735)

Table 2. Percent of true-positive (TP) and false-positive (FP) reports for 0%, 15% and 25% loss.

	0% Loss		15% Loss		25% Loss	
Observer	TP	FP	TP	FP	TP	FP
1	88	16	88	0	80	16
2	80	12	88	16	76	28
3	80	32	84	28	84	40
4	92	24	84	28	80	36
5	64	28	72	32	68	32
6	88	48	92	60	96	64
7	92	36	96	12	84	40
8	60	12	68	4	56	4
9	76	28	76	32	68	32
10	72	12	80	12	72	16
Mean	79	25	83	22	76	31

Table 3. Coefficients of correlation for the number of reported microcalcifications per cluster vs the actual number of microcalcifications for 0%, 15% and 25% loss conditions.

Observer	0% Loss	15% Loss	25% Loss
1	.97	.84	.91
2	.81	.81	.84
3	.79	.76	.79
4	.79	.83	.85
5	.85	.74	.86
6	.87	.83	.88
7	.82	.77	.83
8	.81	.79	.73
9	.86	.79	.92
10	.91	.77	.84

Figure 1.

FIGURE CAPTIONS

Figure 1. Example of a 25% loss image. The image on the left (a) shows a mammographic image with 25% loss before four-point interpolation recovery. Each diagonal line represents a lost packet. The figure on the right (b) shows the same image with 25% loss after four-point interpolation recovery.