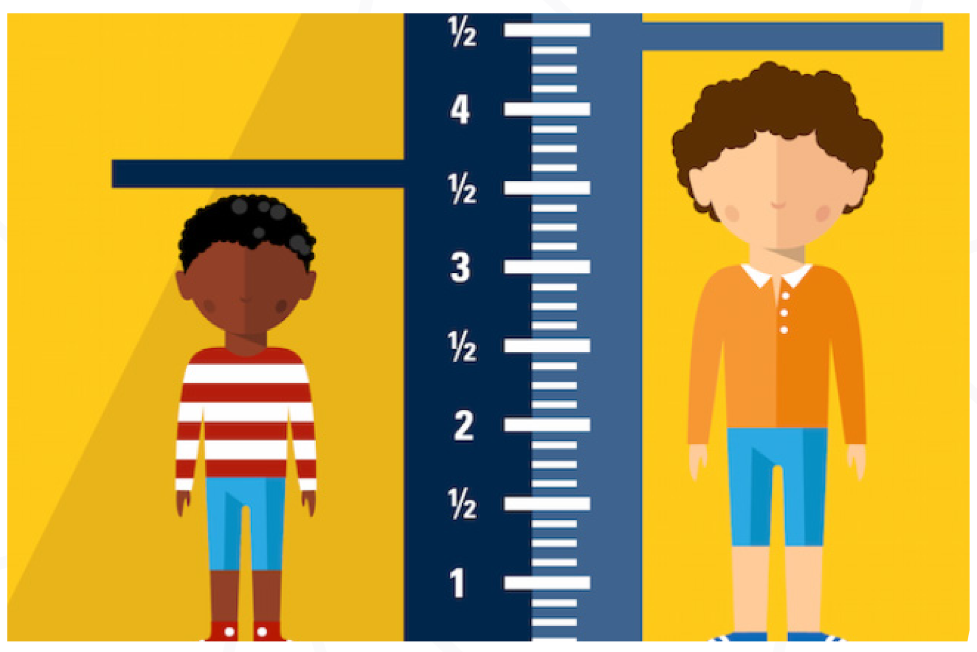
ASSESSING THE IMPACT OF THE DNLM AS A PREPROCESSING STAGE IN A CNN BASED APPROACH FOR AGE ESTIMATION USING DIGITAL HAND X - RAY IMAGES

S. Calderon[1], F. Fallas[2], M. Zumbado[3], P. N. Tyrrell[4], H. Stark[5], Z. Emersic[6], B. Meden[7], M. Solis[8]

Instituto Tecnologico de Costa Rica[1,3,8], Universidad de Costa Rica[2], University of Toronto[4,5], Univerza v Ljubljani[6,7]

PROBLEM CONTEXT

In pediatrics and endocrinology, estimation of skeletal maturity using X-ray images is often performed by physicians interested in comparing patient bone age with their chronological age. The radiological examination analyzes the left hand X-ray image using either the Greulich and Pyle or the Tanner-Whitehouse methods. Such comparisons help to diagnose and observe the effects of endocrine and metabolic disorders. The usage of machine learning to estimate bone age using digital X-ray images has been explored in [1, 2], with recent deep learning based techniques being the most successful approaches.



labblog.uofmhealth.org

PREPROCESSING

Given the major paradigm shift implemented in CNN models evaluating the impact of preprocessing techniques becomes appealing. In [3, 4], the impact of noise and denoising techniques is analyzed using a CNN based approach where [4] shows that popular CNN models like VGG-16 and GoogLeNet, are all negatively susceptible to noise and blur artificial degradations of training samples. In [3] the authors evaluate the impact of implementing NLM denoising for a CNN using the MNIST dataset with gaussian noise. No parameter assessment s shown in such work. Assessing the impact of edge and contrast enhancement and denoising of training samples using for instance the Deceived Non Local Means (DNLM-IFFT) filter [5] which performs both tasks, in a real world application application: the estimation of bone age using digital X-ray images, we consider is important to address.

$$F_{\text{USM}} = U + \lambda L$$

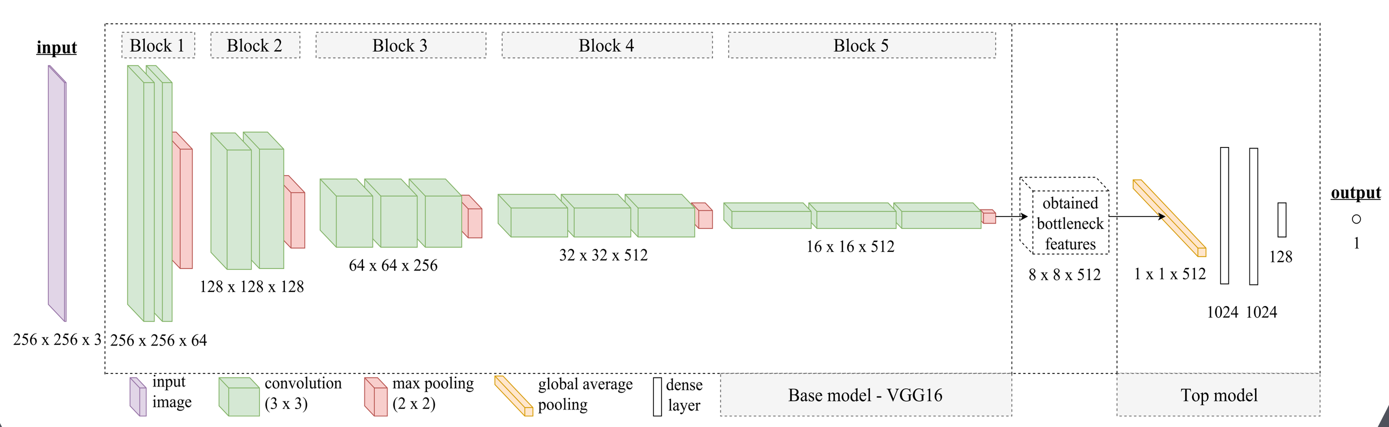
$$Y(p) = \left(\sum_{m \in \Omega} \psi_{\text{NLM}}(U,p,m) \right)^{-1} \left(\sum_{m \in \Omega} \psi_{\text{NLM}}(U,p,m)\, F_{\text{USM}} \right)$$

$$\psi_{\text{NLM}}(U,p,m) = \exp\left(- \frac{\| \vec{\eta}(\omega,m) - \vec{\eta}(\omega,p) \|^2}{h} \right)$$

U is the input image, *L* the laplacian mask, ω the neighbroghood size, λ and h the unsharp gain and blurring parameteres.

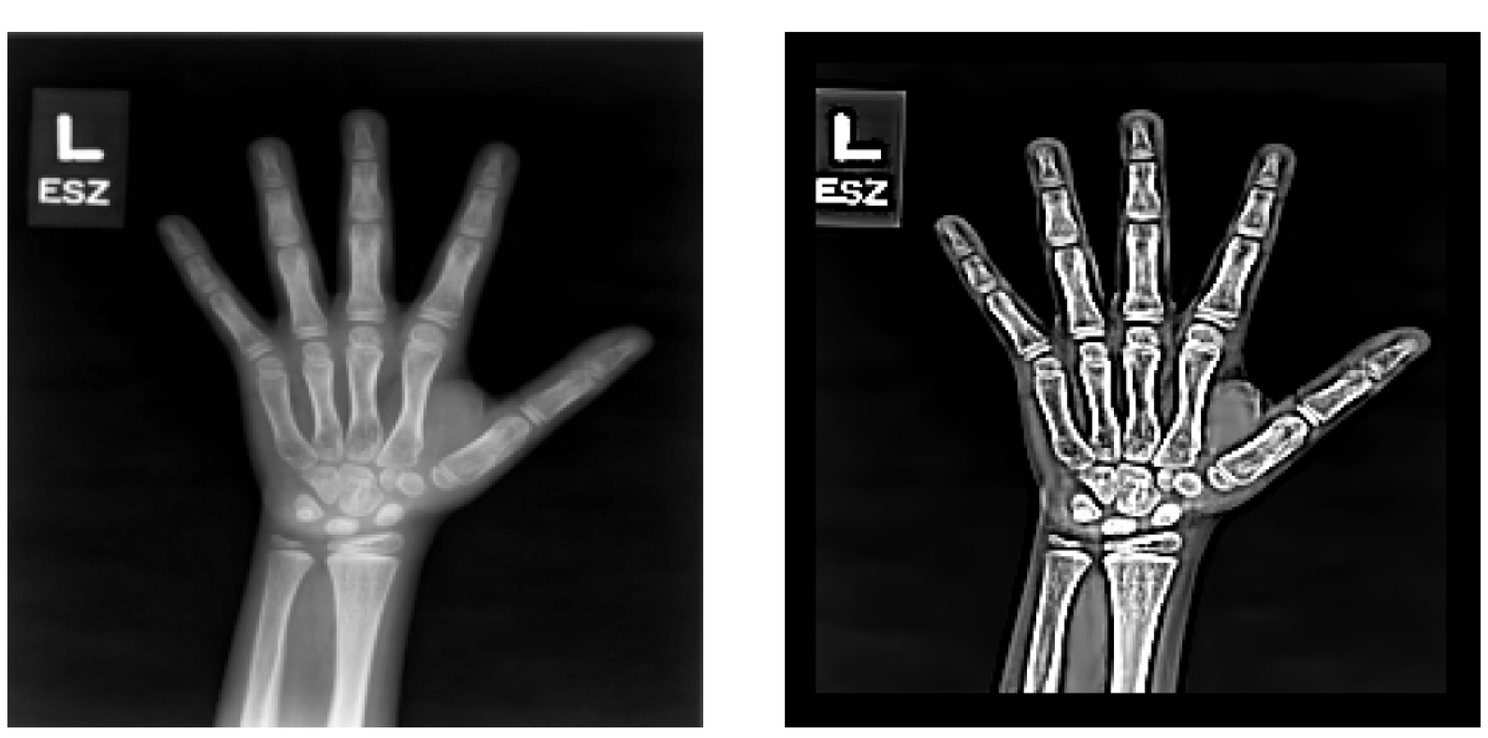
PROPOSED SOLUTION

Implement the the DNLM-IFFT filter, as a preprocessing stage in a common VGG16 model for bone age estimation using digital X-ray images, and assess the impact of λ and h parameters with statistical significance.



DATA AND EXPERIMENTS

The dataset consists in a publicly available RSNA digital images repository from left hands of both male and female subjects, with ages ranging from 1 month to 19 years. Data was acquired from Stanford Children's Hospital and Colorado Children's Hospital. Since λ and h are the most important parameters to be tuned for the DNLM filter, ANOVA was performed to compare performance with statistical significance of tuning them. For both parameters we defined three and four levels respectively, with 12 combinations or treatments. We executed 10 replicas per treatment for a total of 120 runs.



DNLM output with λ = *5*, h = *5*, window size of *15 x 15* and a neighborhood size of *3 x 3*.

RESULTS AND CONCLUSIONS

Our experiment results suggested that image preprocessing for a CNN based approach brings an impressive accuracy boost yielding very similar results to recent previous work, with the lowest MAE of 0.79 [1] and our model achieving a similar MAE of 0.79 years, with a larger dataset. The ANOVA showed that incrementing both parameters λ and h postively impact the CNN performance with statistical significance, yielding a 42% accuracy boost over the base line model. As future work, given the concluded importance of preprocessing for CNN models, we aim to work in CNN architectures which implement preprocessing approaches, calibrating its parameters along the model [6]. Table 1 shows a descriptive analysis of the results.

λ	h	**Avg.**	**Std.**	**Min.**
0	0	22.203	1.033	20.989
0	8	20.457	1.39	18.168
0	12	19.371	1.399	17.801
0	14	18.777	0.519	18.137
2.5	0	22.367	0.961	20.699
2.5	8	16.166	1.343	14.058
2.5	12	14.343	1.404	12.866
2.5	14	14.998	0.837	13.42
5	0	21.079	0.853	20.11
5	8	12.956	1.201	11.852
5	12	**12.803**	0.993	11.864
5	14	**12.889**	0.713	11.56

Table 1: RMSE Descriptive analysis with different parameter values for the DNLM-IFFT.

REFERENCES

[1] C. Spampinato, S. Palazzo, D. Giordano, M. Aldinucci, and R. Leonardi, "Deep learning for automated skeletal bone age assessment in x-ray images," Medical image analysis, vol. 36, pp. 41–51, 2017.

[2] H. Lee, S. Tajmir, J. Lee, M. Zissen, B. A. Yeshiwas, T. K. Alkasab, G. Choy, and S. Do, "Fully automated deep learning system for bone age assessment," Journal of Digital Imaging, pp. 1–15, 2017.

[3] T. Nazare, G. P. da Costa, W. Contato, and M. Ponti, "Deep convolutional neural networks and noisy images," in Iberoamerican Conference on Pattern Recognition (CIARP), 2017.

[4] S. Dodge and L. Karam, "Understanding how image quality affects deep neural networks," in Quality of Multimedia Experience (QoMEX), 2016 Eighth International Conference on, pp. 1–6, IEEE, 2016.I.

[5] S. Calderon and M. Zumbado, "Dnlm-IIFFT: An implementation of the deceived non local means filter using integral images and the fast fourier transform for a reduced computational cost," in Iberoamerican Conference on Pattern Recognition (CIARP), 2017.

[6] V. Jampani, M. Kiefel, P. Gehler. "Learning sparse high dimensional filters: Image filtering, dense CRFS and bilateral neural networks." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016.

Parma Group

https://github.com/PARMAGroup/BoneAgeAssessment