Skin cancer is the most common cancer worldwide, and both non-melanoma and melanoma have been growing in recent years [1]. According to the World Health Organization, there are 132,000 skin cancers with melanoma and between 2 and 3 million non-melanoma skin cancers worldwide annually [2]. Given the fact that melanoma accounts for just around 1% of skin cancers, it nevertheless causes a significant number of deaths from skin cancer. The risk of melanoma rises with rising age and is more common in men than in women. According to the latest statistics, around 100,350 new melanomas are expected to be diagnosed, and about 6,850 people in the United States will die of melanoma by 2020[3]. While melanoma is a malignant type of skin cancer, good types such as seborrheic keratosis and nevi also exist. [4], [5].

The 5-year survival rate for melanoma is 99 percent when melanoma is diagnosed early and increases the prognosis of patients with malignant melanoma [1],[6]. Melanoma diagnosis actually calls for early and reliable detection by experts. Dermoscopy is a non-invasive procedure to assist the dermatologists with detection, which allows greater detection precision than is possible with the naked eye. . It is a method of high-resolution imaging in which reflection of the skin is reduced, making the deeper underlying structures in the skin more visible [7]. When determining whether a skin lesion is benign or malignant, the ABCD rule is applied by default in dermoscopic images where asymmetry (A), border (B), color (C) and diameter (D) are assessed [8].

Although the accuracy of experts in assessing dermoscopic images is between 75% - 84%, it requires considerable experience depending on the expert [7]. Nevertheless, manual assessment is particularly time consuming and prone to error, especially when the difference between a benign and malignant skin lesion is nil [9]. A promising tool that has emerged in recent years to assist dermatologists in assessing dermoscopic images is deep learning. Recent studies show that deep learning outperforms dermatologists in skin lesion classification [10], [11], [12]. While deep learning is a promising tool for classifying skin lesions, it also has challenges, since deep neural networks require very large training sets [13]. To circumvent that difficulty, "transfer learning" is implemented, a technique in which a trained base network is transferred to a second target network and then trained on the target dataset and task [14].

Transfer learning is currently not widely used in the medical context, as there is a shortage of annotated data, and it is often expensive [15]. Crowdsourcing is implemented to tackle this issue, in which crowds of staff are asked to label images. Recently, this has become increasingly apparent in the field of deep learning, since labels cannot be produced by using automated methods. Additionally, many research questions do not have labels [16]. Since crowds are unable to assess dermoscopic images of skin lesions for benign or malignant when much experience is required, it is suggested that the crowds look at the visual characteristics such as shape and color. The combination of the annotated data and transfer learning is called "multi-task learning" where the data yield promising benefits [17].

This research builds on the work of Raumanns et al. In that work, VGG-16 is used for training the multi-task model. VGG-16 is a convolutional neural network and is characterized by its simplicity, it only uses 3x3 convolutional layers stacked on top of each other. Similar to the VGG-16 network is the VGG-19 network where the network has 19 weight layers instead of 16. A disadvantage of the VGG networks is that they often have a large architecture [18]. Although these networks are often used with many image classification problems, it is also desirable to use networks with narrow network architectures, an example of such a network is the InceptionV3 network. This network is made by Google and is deeper and more complex than the VGG networks but has the advantage that the computational cost is much lower than the VGG networks [19], [20]. Since the work of Raumanns et al. Is based on the VGG-16 network, it is also interesting to look at the impact of other networks on multi-task learning. So, the goal of this study is to compare the existing baseline model (VGG-16) with the VGG-19 and the InceptionV3 model to see what impact it has on the performance of multi-task learning.

According to literature, classification of skin lesions in a dilated InceptionV3 network achieves higher accuracy than in the dilated VGG-16 and dilated VGG-19 network model. However, the dilated InceptionV3 and dilated VGG-16 network do not differ much from each other, 90.95% versus 90.10% respectively [21]. In another study looking at breast cancer histology images classification, the VGG-16 compared to the VGG-19 network had higher accuracy and Area Under Curve (AUC) using both Fine-tuned Pre-trained Network and Full-trained Network. However, the networks did not differ much from each other in results, but only a few percentages. [22]. In yet another study, it was claimed that when classifying melanoma dermoscopic images, the InceptionV3 model works best in contrast to the VGG-16 and VGG-19 model where the InceptionV3 model had an average accuracy of 93.7% and the VGG-16 and VGG-19 model 74.3% and 76.6% respectively [23]. Since the literature indicates that the InceptionV3 network model performs better than the VGG networks, multi-task learning is expected to perform better with the InceptionV3 network model. Moreover, multi-task learning is also expected to do better on the VGG-19 network model than the VGG-16 network model.

[1] Skin Cancer Facts & Statistics. (n.d.). Retrieved from https://www.skincancer.org/skin-cancer-information/skin-cancer-facts/

[2] Skin cancers. (2017, October 16). Retrieved from https://www.who.int/uv/faq/skincancer/en/index1.html

[3] Information and Resources about for Cancer: Breast, Colon, Lung, Prostate, Skin. (n.d.). Retrieved from https://www.cancer.org/

[4] Wollina U. (2018). Seborrheic Keratoses - The Most Common Benign Skin Tumor of Humans. Clinical presentation and an update on pathogenesis and treatment options. *Open access Macedonian journal of medical sciences*, *6*(11), 2270–2275. https://doi.org/10.3889/oamjms.2018.460

[5] Fujisawa, Y., Inoue, S., & Nakamura, Y. (2019). The Possibility of Deep Learning-Based, Computer-Aided Skin Tumor Classifiers. *Frontiers in medicine*, *6*, 191. https://doi.org/10.3389/fmed.2019.00191

[6] Jerant, A. F., Johnson, J. T., Sheridan, C. D., & Caffrey, T. J. (2000, July 15). Early Detection and Treatment of Skin Cancer. Retrieved from https://www.aafp.org/afp/2000/0715/p357.html

[7] Ali, A.-R. A., & Deserno, T. M. (2012). A systematic review of automated melanoma detection in dermatoscopic images and its ground truth data. *Medical Imaging 2012: Image Perception, Observer Performance, and Technology Assessment*. doi: 10.1117/12.912389

[8] Stolz, W. (1995). The ABCD rule of dermatoscopy: high negative predictive value for the recognition of malignant melanomas. *Journal of the European Academy of Dermatology and Venereology*, *5*(1). doi: 10.1016/0926-9959(95)95977-9

[9] R D, S., & A, S. (2019). Deep Learning Based Skin Lesion Segmentation and Classification of Melanoma Using Support Vector Machine (SVM). *Asian Pacific journal of cancer prevention : APJCP*, *20*(5), 1555–1561. https://doi.org/10.31557/APJCP.2019.20.5.1555

[10] Brinker, T. J., Hekler, A., Enk, A. H., Berking, C., Haferkamp, S., Hauschild, A., … Utikal, J. S. (2019). Deep neural networks are superior to dermatologists in melanoma image classification. *European Journal of Cancer*, *119*, 11–17. doi: 10.1016/j.ejca.2019.05.023

[11] Brinker, T. J., Hekler, A., Enk, A. H., Klode, J., Hauschild, A., Berking, C., … von Kalle, C. (2019). Deep learning outperformed 136 of 157 dermatologists in a head-to-head dermoscopic melanoma image classification task. *European Journal of Cancer*, *113*, 47-54. https://doi.org/10.1016/j.ejca.2019.04.001

[12] Hekler, A., Utikal, J. S., Enk, A. H., Solass, W., Schmitt, M., Klode, J., … Brinker, T. J. (2019). Deep learning outperformed 11 pathologists in the classification of histopathological melanoma images. *European Journal of Cancer*, *118*, 91–96. doi: 10.1016/j.ejca.2019.06.012

[13] Shin, H. C., Roth, H. R., Gao, M., Lu, L., Xu, Z., Nogues, I., Yao, J., Mollura, D., & Summers, R. M. (2016). Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. *IEEE transactions on medical imaging*, *35*(5), 1285–1298. https://doi.org/10.1109/TMI.2016.2528162

[14] Yosinki, J., Clune, J., Bengio, Y., & Lipson, H. (2014). How transferable are features in deep neural networks? *Proceedings of the 27th International Conference on Neural Information Processing Systems*, *2*, 3320–3328.

[15] Thaler, S., & Menkovski, V. (2019). The Role of Deep Learning in Improving Healthcare. *Data Science for Healthcare*, 75–116. doi: 10.1007/978-3-030-05249-2_3

[16] Ye, C., Coco, J., Epishova, A., Hajaj, C., Bogardus, H., Novak, L., Denny, J., Vorobeychik, Y., Lasko, T., Malin, B., & Fabbri, D. (2018). A Crowdsourcing Framework for Medical Data Sets. *AMIA Joint Summits on Translational Science proceedings. AMIA Joint Summits on Translational Science*, *2017*, 273–280.

[17] Raumanns, R., Contar, E.K., Schouten, G & Cheplygina, V. (2020). Multi-task Learning with Crowdsourced Features Improves Skin Lesion Diagnosis (cite arXiv:2004.14745)

[18] Simonyan, K. & Zisserman, A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition (cite arXiv:1409.1556)

[19] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J. & Wojna, Z. (2015). Rethinking the Inception Architecture for Computer Vision (cite arxiv:1512.00567)

[20] Grm, K., Struc, V., Artiges, A., Caron, M. & Ekenel, K. (2017). Strengths and Weaknesses of Deep Learning Models for Face Recognition Against Image Degradations (cite arXiv:1710.01494)

[21] Ratul, A. R., Mozaffari, M. H., Lee, W.-S., & Parimbelli, E. (2019). Skin Lesions Classification Using Deep Learning Based on Dilated Convolution. doi: 10.1101/860700

[22] Shallu, & Mehra, R. (2018). Breast cancer histology images classification: Training from scratch or transfer learning? *ICT Express*, *4*(4), 247–254. doi: 10.1016/j.icte.2018.10.007

[23] Cui, X., Wei, R., Gong, L., Qi, R., Zhao, Z., Chen, H., … Gao, X. (2019). Assessing the effectiveness of artificial intelligence methods for melanoma: A retrospective review. *Journal of the American Academy of Dermatology*, *81*(5), 1176–1180. doi: 10.1016/j.jaad.2019.06.042