DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
INDIAN INSTITUTE OF INFORMATION TECHNOLOGY, DESIGN AND MANUFACTURING KANCHEEPURAM
CHENNAI - 600127

Synopsis Of

Machine Learning Algorithms for Abnormal Entity Detection

A Thesis

To be submitted by

JOSHI PRATIK

For the award of the degree

Of

DOCTOR OF PHILOSOPHY

1 Abstract

Abnormal entity detection is the process of finding patterns in data that do not follow normal patterns of behavior. False data points or abnormal entities that significantly distort the sample or push the boundaries of the data set are rare. In some cases, abnormal entities suggest that the pattern is too narrow and that the boundaries need to be corrected. In the literature, domain specific abnormal entities detection have been explored. But when the data changes from domain to domain, different set of challenges come across. This thesis explores different domains with varying data and provides different machine learning approaches for detecting the abnormal entities in those respective domains. The domains explored are pharmocovigilance, medical images and facial images and organized based on the input data.

The issue of having fewer reports makes it more challenging to predict ADRs. The problem of under-reporting of adverse events exists worldwide, which makes it difficult to develop objective predictive models. In this thesis, we proposed a new method for predicting adverse events by balancing data sets. A core-optimized SVM was trained using each balanced data set, and a certain ADR prediction for a particular drug was obtained by voting for the combined optimal SVM. Despite good performance, the proposed model requires separate model per ADR which takes significant space and complexity. To tackle this, another method is developed which takes one model to predict any ADR and also gives improved performance. We developed and trained a custom deep neural network (DNN) called the Knowledge Graph DNN (KGDNN) to predict ADR using KG embedding. Using these incorporations, ADRs are classified according to the KGDNN model. To demonstrate the effectiveness of the model, two case studies of liver-damaging drugs and drugs prescribed for COVID-19 were performed.

In this thesis, two methods have been proposed to locate the optical disc using segmentation performed on an interference map derived from a group of generalized motion patterns in the retinal fundus image. Also, a few studies have been performed to classify a particular retinal image into a normal fundus image and an abnormal fundus image. This thesis presents a new method of transfer learning for the detection of fundus anomalies. EfficientNetV2 was used as a classification model to classify normal and abnormal images of the fundus. The proposed model is tested against modern models, including transformers and models based on MultiLayer Perceptron (MLP), which work well in image classification tasks. It has been observed that convolutional neural networks still outperform transformer-based and MLP-based models in detecting abnormal fundus images.

We live in a digital age where information travels very quickly. Images and videos make up the majority of the material. However, the validity of such data is under doubt. Particularly, the DeepFakes threat hangs over digital media in the way of revenge porn, necessitating the development of effective systems to combat it. We propose a model to detect obscene conten with EfficientNet B4 as a base model. The current facial detection methods frequently fall short when real-world data is released because they are biased toward certain datasets. In this study, we propose an novel, comprehensible, and effective method for identifying facial forgery called GaborEffNet. The proposed method is found to be effective in generalisation after performing cross-dataset testing.

2 Objectives

Objectives of the thesis are as follows

- To develop algorithms for predicting adverse drug reactions
  - To handle the imbalance data and find an efficient way to balance for predicting adverse drug reactions
  - To build a single and efficient model for predicting adverse drug reactions
- To develop algorithms on retinal fundus images
  - To develop methods to segment the optic disc
  - To detecting the abnormal fundus images from normal fundus images
- To develop algorithms for forged obscene content detection
  - To develop a model for detecting obscene content in images/videos
  - To develop a model for detecting facial forgery in images/videos

3 Existing Gaps Which Were Bridged

This thesis presents predictive models for abnormal entity detection across different domains. The domains explored in particular are pharmacovigilance, retinal fundus images and facial images. There is a scope in improving the performances and developing better models in each domain.

- **Pharmacovigilance**
  - The absence of information regarding the ADRs reported is one of the major issues with ADR prediction. As a result, the SIDER database has only recorded a very small number of positive samples(Kuhn *et al.*, 2016). As a result, prediction models are skewed in favour of negative samples. Therefore, a novel method has been proposed to nullify or lessen the dataset's skewness.
  - Except for the model proposed by Zhang *et al.* (2021), all other models either utilise single model for each ADR or test on a small number of medications and ADRs Despite the use of biological and phenotypic features by Jamal *et al.* (2017), there are very few drugs and ADRs. The knowledge graphs have been employed by Bean *et al.* (2017) and Zhang *et al.* (2021) to examine the ADRs, although their knowledge graph is restricted to only ADRs, indications, and target proteins. Additionally, only 10 ADRs were used in the Bean *et al.* (2017) experiments.
- **Retinal fundus images**
  - Although many segmentation approaches have been proposed to segment the optic disc(Priyadharsini *et al.*, 2018; Khalil *et al.*, 2017), there is huge scope in improving the metrics. Particularly, unsupervised methods lack the performance and it is of high priority since there is limited data available in medical domain.

  - To help eye specialists and to identify diseases in their early stages, automatic disease diagnosis is necessary. Using fundus imaging systems, which effectively record the abnormalities inside the eye, can be an excellent way to find disorders. When there are little medical resources available, mass screening of the population can also benefit from the automatic detection of diseases. The current research focuses on detecting a particular type of disease but the approaches to classify abnormal and normal fundus images are limited.

- **Facial images**
  In our nation, women are frequently harassed using fake pornographic photographs, and there is no effective system in place to stop this. Even though our nation has severe cyber laws, only a small portion of offenders are apprehended by the police. To combat this, a framework has been put up to stop the uploading of fake porn movies to social media platforms.
  - Pornography detection is a difficult challenge since the films feature people with a variety of skin tones, including Asian, White, and Multi-Ethnic people, as well as persons with a wide range of skin tones, from White to Brown to Dark. The challenging non-pornographic category in the NPDI dataset presents additional difficulties. However, the dataset is deficient in videos shot from various angles and non-consensual videos that may be found online.
  - The state-of-the-art techniques currently have difficulty in detecting facial forgeries in photographs and films that are freely available online(Chollet, 2017; Afchar *et al.*, 2018; Zi *et al.*, 2020). Although the proposed solution outperforms the current systems, there is still much room for improvement.

4 Most Important Contributions

4.1 Prediction of Adverse Drug Reactions

4.1.1 An ensembled SVM based approach

By altering the negative samples and fixing the positive samples, the dataset has been divided into numerous datasets with balanced data. Equal numbers of positive and negative samples are present in each partition. ADRs are caused by the prescribed medicine, according to the positive sample. Here, we've treated drugs that don't induce ADRs and ADRs that haven't been reported as negative samples. 80% of the samples in each dataset are taken into account for training, while 20% are taken into account for testing. Samples from each dataset are used to train the Support Vector Machine (SVM) model as given in the Fig. 1. Following training, each SVM model was tested using a single file that contained all the testing data that had been isolated from each dataset. Based on both soft voting and hard voting, the outcome was determined as given in the Fig. 2. For every ADR in the database, the aforementioned procedure is repeated. It has been observed that the proposed model has outperformed the existing methods.

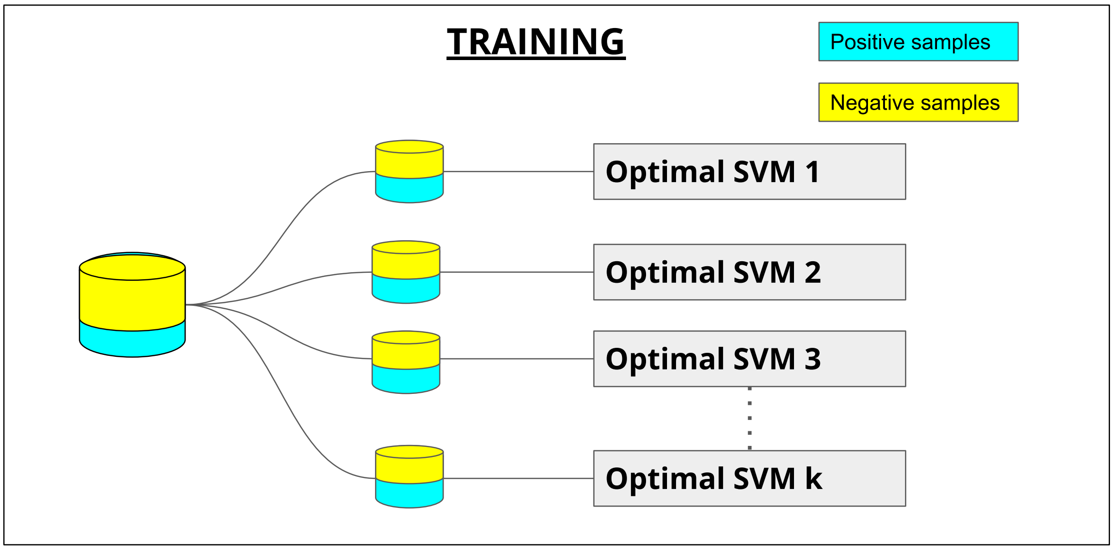

Figure 1: Training phase of the method

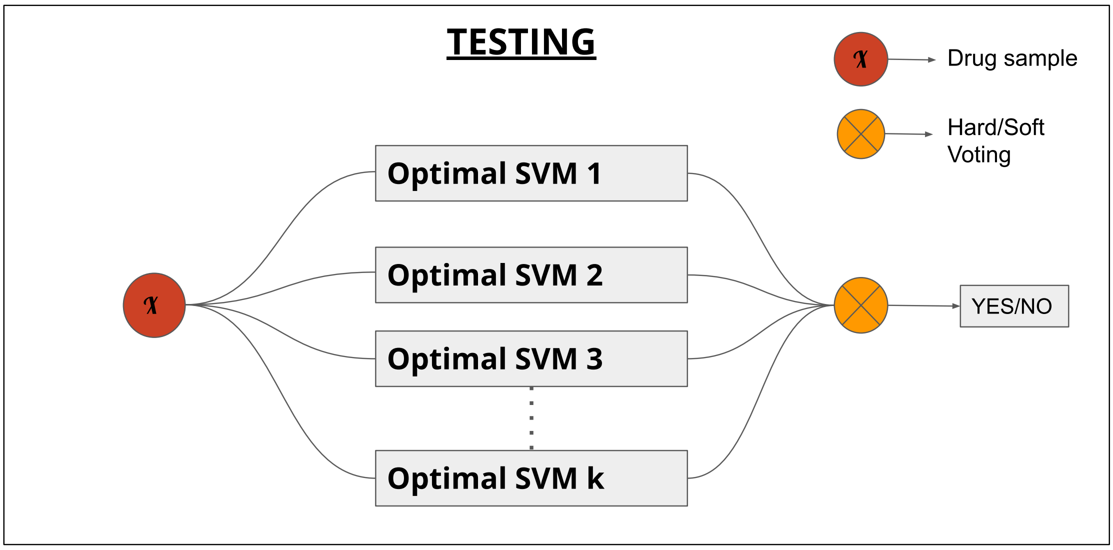

Figure 2: Testing phase of the method

4.1.2 A knowledge graph embedding based approach

DNN has so far been employed in the literature in a different approach to find ADR. The association between two nodes has never been utilised as a labelled example in a DNN. In the literature, DNN is found to be an efficient classification model when the number of examples is respectably high. In order to comprehend the association between the nodes u and v where $u \neq v$ and $u, v \in D$ where $D = \{$drugs, ADRs, indications, target proteins, genes, pathways$\}$, a custom-made DNN has been proposed. Using the context data on the relationship between the drug-target protein, drug-pathway, and drug-gene interaction, one may predict the ADRs connected to drugs. Additionally, it has been documented in numerous clinical research that the adverse medication reaction is dependent on pathways and gene information(Zheng *et al.*, 2014), which is why we added these two types of nodes (genes and pathways) to the proposed model. In light of this, it is anticipated that our prediction model will perform better if it incorporates this additional data in addition to the more common target protein, indication, etc. Additionally, we have confirmed this through experiments. The overview of the proposed method is given in Fig. 3. It has been found to be advantageous to add new features and to create a customised DNN model. The SIDER database has been used to examine the proposed method, which produced improved results when compared to existing systems.

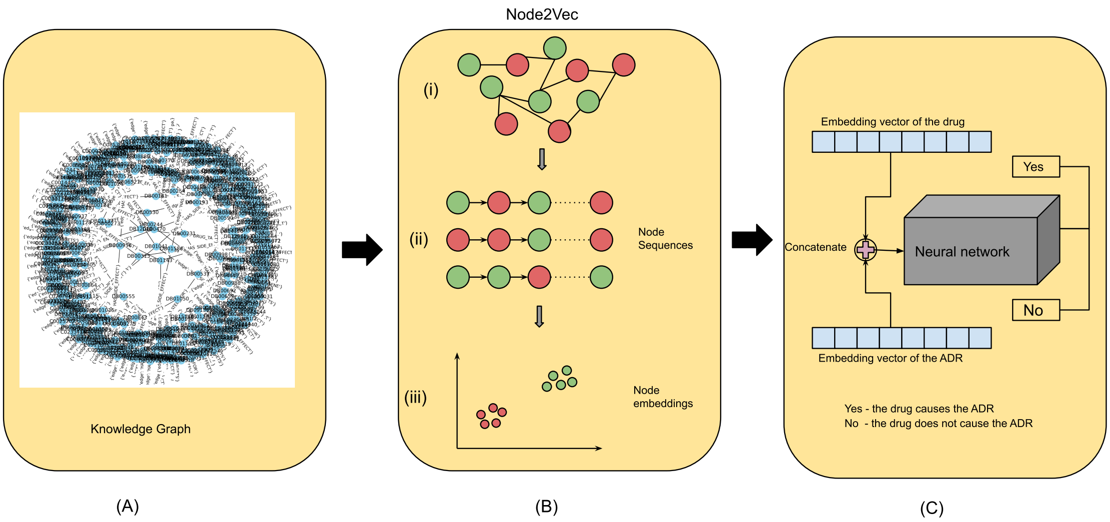

Figure 3: Overview of the proposed method A) The knowledge graph constructed using drugs, ADRs, indications, targets, pathways and genes B) (i) The constructed knowledge graph is passed as input to the Node2Vec algorithm (ii) Node sequences are generated by Node2Vec algorithm using random walk(iii) The node sequences are given as input to the Word2Vec algorithm and the node embeddings are obtained (C) The embeddings of drug and ADR are concatenated and given as input feature vector to a deep neural network which acts as prediction model for ADR classification

4.2 Optic Disc Segmentation and Abnormal Fundus Images Detection

- Two methods of segmenting optical discs have been introduced. Applying the movement of the subject in the fundus improves the contrast between the optical disc and the background. The proposed method 1 performs better in terms of efficiency and it is faster than the competitive method, but suffers from poor performance for dimly lit images. The proposed method 2 uses iterative graph cutting, outperforms the competing method on the Drishti-GS1 reference data set, and is also found to be robust to Gaussian noise. The pipeline of the proposed methods is given in the Fig. 4.

- A novel method based on transfer learning is proposed to distinguish between normal and abnormal fundus images. As a starting point, the EfficientNetv2(Tan and Le, 2021) architecture was employed. Recently developed image classification methods, such as transformer-based models(d'Ascoli *et al.*, 2021; Xu *et al.*,

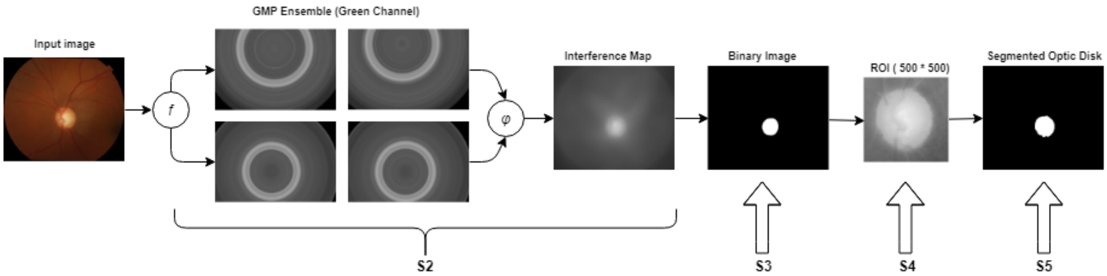

Figure 4: Pipeline of the proposed approaches

2021) and MLP-based models(Tolstikhin *et al.*, 2021), have been studied. Experiments have been conducted using a private dataset that was obtained from a hospital and manually labelled by eye specialists. Our dataset has been used to implement and test the current methodologies. The outcomes demonstrated that transfer learning on EfficientNetV2 outperformed the conventional approaches.

4.3 Forged Obscene Content Detection

- For detecting the obscene content, the EfficientNet B4(Tan and Le, 2019) model has been used as base model. An EfficientNet inspired model has been proposed using EfficientNet and a mean pooling and maximum pooling. We concatenated the results of average pooling and maximum pooling and propagated to dense layers of size 512, 128. The proposed model has been evaluated on a benchmark dataset called NPDI. It has been observed that the proposed model has outperformed the existing methods on NPDI dataset.

- A novel deep learning model, named GaborEffNet has been proposed. This model incorporates learnable Gabor filter, EfficientNet and Channel Attention to detect facial forgery. The proposed method uses learnable Gabor filters to differentiate the fake faces from the real ones. A traditional Gabor filter has been incorporated into the architecture of a Convolutional Neural Network(CNN) as first layer of the network. It has been observed that Gabor filter when applied to the forged face is able to discriminate from the real face as shown in the Fig. 5. In the Fig. 5, the real source and fake images are taken from UADFV dataset and the target image is taken from the internet. In the bottom row, the Gabor filters applied on source, target and fake are given, and at the end the Gabor filter is given. From the figure, it can be observed that the Gabor filter when applied on fake image clearly differentiates from the real ones. Also, a channel attention module is also incorporated into the network which aids in helping to find the manipulations across the channels. The proposed model outperformed existing methods on benchmark datasets.

5 Conclusions

Detecting abnormal entities in different domains have different set of challenges. Machine learning algorithms have demonstrated success in a number of domains. Us-

Figure 5: Gabor filter applied on fake and real faces

ing machine learning algorithms for detecting the abnormal entities can aid in solving various problems. This thesis explores developing different machine learning based approaches for detecting abnormal entities in different domains. The domains that are explored in particular are pharmacovigilance(adverse drug reactions), facial images(forged obscene content detection), and medical images(retinal fundus images). In a brief, the work done can be concluded in following way

- Algorithms to detect the adverse drug reactions are proposed. In particular, two algorithms have been developed. The first one uses the chemical properties of the drug as the feature vector. This method discusses a novel way of dealing with unbalanced datasets by partitioning the dataset such that many datasets with equal number of positive and negative samples are created. The method then uses many optimal Support Vector Machines(SVM) to make inferences and decision. The second method is based on knowledge graph embedding. The feature vector is obtained using the node embeddings of the graph. A Deep Neural Network(DNN) has been designed to train and test the model.

- Algorithm to detect the abnormal fundus images is discussed along with an algorithm to detect the optic disc in fundus images. To detect the abnormal fundus images, a transfer learning based approach has been developed. Various state-of-the-art algorithms have been investigated and observed the most effective model for transfer learning on fundus images. In the work, an algorithm to localize and segment the optic disc has been discussed. The Generalized Motion Patterns(GMP) are generated for the fundus images. Inference maps are obtained from those GMPs. Later, optic disc has been segmented using thresholding on those inference maps. Also, in another work, those segmented images have been better segmented using GrabCut algorithm.

- Algorithms to predict forged obscene content are discussed. For detecting the obscene content, the EfficientNet B4 inspired model has been used as base model.

We got the output from EfficientNet inspired model and did a mean pool and maximum pooling. We concatenated the results of average pooling and maximum pooling and propagated to dense layers of size 512, 128. The proposed model has been evaluated on a benchmark dataset called NPDI. It has been observed that the proposed model has outperformed the existing methods on NPDI dataset. Although the model performed well for easy porn and easy non porn videos, the model struggles to detect hard porn and hard non porn videos. For detecting abnormalities in facial images a novel method to detect facial forgeries is discussed. An observation has been made that the Gabor filter helps in discriminating the real faces from fake faces. Therefore, using Gabor filter a deep learning model called Gabor EfficientNet has been designed and developed to detect the facial forgeries.

6 Organization of the Thesis

The chapters are organized based on the type of data. Chapter 1 uses 1D data which are drugs in the form of SMILE strings to detect abnormalities, chapter 2 uses 2D images which are fundus images to detect the abnormalities, and chapter 3 uses 2D images and 3D videos to detect forgery and obscenity. The overview of the thesis is given in the Fig. 6. The thesis further organized into the following chapters: The proposed outline

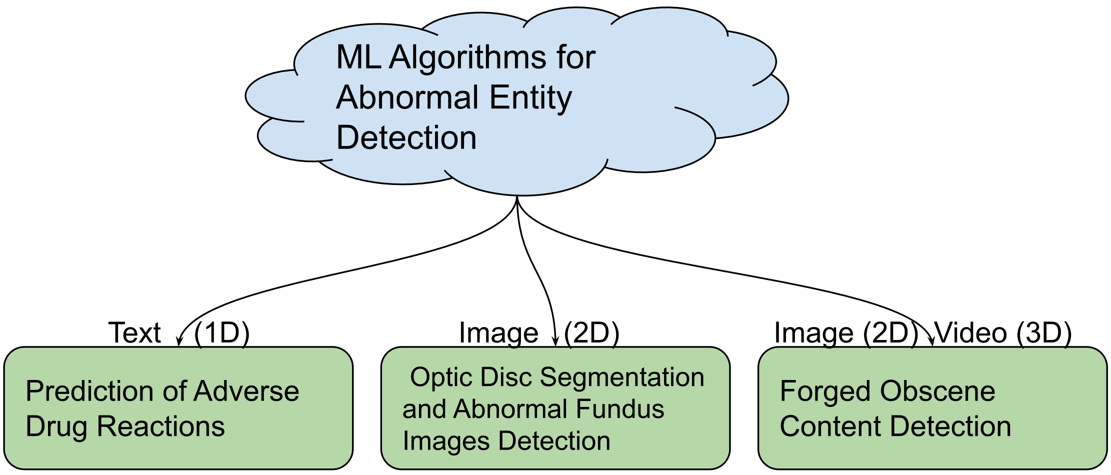

Figure 6: Overview of the work

of the thesis is as follows:

(a) Chapter 1: Introduction

(b) Chapter 2: Literature Survey

(c) Chapter 3: Prediction of Adverse Drug Reactions

(d) Chapter 4: Optic Disc Segmentation and Abnormal Fundus Images Detection

(e) Chapter 5: Forged Obscene Content Detection

(f) Chapter 6: Conclusion and Future Scope

7 List of Publications

Journals

(a) P. Joshi, V. Masilamani, and A. Mukherjee, "A knowledge graph embedding based approach to predict the adverse drug reactions using a deep neural network", *Journal of Biomedical Informatics*, 132, pp. 104122, 2022.

(b) P. Joshi, V. Masilamani, and R. Ramesh, "An ensembled svm based approach for predicting adverse drug reactions", *Current Bioinformatics*, 16(3), pp. $422-432$, 2021.

(c) P. Joshi, R. Raj KS, V. Masilamani and J. Alike, "Optic disc localization using interference map and localized segmentation using grab cut", *Automatika*, 62(2), pp. $187-196$, 2021.

Conferences

(a) P. Joshi, and V. Masilamani, "An Efficient Transfer Learning Based Approach for Detecting the Abnormal Fundus Images." *In 2021 5th Conference on Information and Communication Technology (CICT)*, pp. $1-5$. IEEE, 2021.

(b) P. Joshi, Ranjth Raj KS, V. Masilamani, Jahnavi Alike, K. Suresh, and K. Kumaresh, "Optic disc localization using interference map and localized segmentation." *In 2019 IEEE 1st International Conference on Energy, Systems and Information Processing (ICESIP)*, pp. $1-4$. IEEE, 2019.

Other publications (not included in thesis)

(a) A. S. Mantripragada, S. P. Teja, R. R. Katasani, P. Joshi, V. Masilamani, and R. Ramesh, "Prediction of adverse drug reactions using drug convolutional neural networks", *Journal of Bioinformatics and Computational Biology*, 19(01), pp. 2050046, 2021.

(b) S. Kiruthika, V. Masilamani and P. Joshi, " Fusion of image quality assessment and transfer learning for COVID19 detection using CT scan image, " *In Proceedings of the Twelth Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP '21)*, ACM, Article 48, pp. $1-9$, 2021.

References

1. **D. Afchar**, **V. Nozick**, **J. Yamagishi**, and **I. Echizen** (2018). Mesonet: a compact facial video forgery detection network. *2018 IEEE International Workshop on Information Forensics and Security (WIFS)*, 1–7.

2. **D. M. Bean**, **H. Wu**, **E. Iqbal**, **O. Dzahini**, **Z. M. Ibrahim**, **M. Broadbent**, **R. Stewart**, and **R. J. Dobson** (2017). Knowledge graph prediction of unknown adverse drug reactions and validation in electronic health records. *Scientific reports*, **7**(1), 1–11.

3. **F. Chollet** (2017). Xception: Deep learning with depthwise separable convolutions. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1251–1258.

4. **S. d'Ascoli**, **H. Touvron**, **M. Leavitt**, **A. Morcos**, **G. Biroli**, and **L. Sagun** (2021). Convit: Improving vision transformers with soft convolutional inductive biases. *arXiv preprint arXiv:2103.10697*.

5. **S. Jamal**, **S. Goyal**, **A. Shanker**, and **A. Grover** (2017). Predicting neurological adverse drug reactions based on biological, chemical and phenotypic properties of drugs using machine learning models. *Scientific reports*, **7**(1), 1–12.

6. **T. Khalil**, **M. U. Akram**, **S. Khalid**, and **A. Jameel** (2017). Improved automated detection of glaucoma from fundus image using hybrid structural and textural features. *IET Image Processing*, **11**(9), 693–700.

7. **M. Kuhn**, **I. Letunic**, **L. J. Jensen**, and **P. Bork** (2016). The sider database of drugs and side effects. *Nucleic acids research*, **44**(D1), D1075–D1079.

8. **R. Priyadharsini**, **A. Beulah**, and **T. S. Sharmila** (2018). Optic disc and cup segmentation in fundus retinal images using feature detection and morphological techniques. *CURRENT SCIENCE*, **115**(4), 748.

9. **M. Tan** and **Q. Le** (2019). Efficientnet: Rethinking model scaling for convolutional neural networks. *In International Conference on Machine Learning*. PMLR.

10. **M. Tan** and **Q. V. Le** (2021). Efficientnetv2: Smaller models and faster training. *arXiv preprint arXiv:2104.00298*.

11. **I. Tolstikhin**, **N. Houlsby**, **A. Kolesnikov**, **L. Beyer**, **X. Zhai**, **T. Unterthiner**, **J. Yung**, **D. Keysers**, **J. Uszkoreit**, **M. Lucic**, *et al.* (2021). Mlp-mixer: An all-mlp architecture for vision. *arXiv preprint arXiv:2105.01601*.

12. **W. Xu**, **Y. Xu**, **T. Chang**, and **Z. Tu** (2021). Co-scale conv-attentional image transformers. *arXiv preprint arXiv:2104.06399*.

13. **F. Zhang**, **B. Sun**, **X. Diao**, **W. Zhao**, and **T. Shu** (2021). Prediction of adverse drug reactions based on knowledge graph embedding. *BMC Medical Informatics and Decision Making*, **21**(1), 1–11.

14. **H. Zheng**, **H. Wang**, **H. Xu**, **Y. Wu**, **Z. Zhao**, and **F. Azuaje** (2014). Linking biochemical pathways and networks to adverse drug reactions. *IEEE transactions on nanobioscience*, **13**(2), 131–137.

15. **B. Zi**, **M. Chang**, **J. Chen**, **X. Ma**, and **Y.-G. Jiang** (2020). Wilddeepfake: A challenging real-world dataset for deepfake detection, 2382–2390.